PATENT
ATTORNEY DOCKET: 62785.000004

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:	)	Confirmation No. 7580
Richard William Falla LE PAGE et al.	)	Group Art Unit: 1645
Application Number: 09/769,744	)	Examiner: DEVI
Filed: January 26, 2001	)	

For: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS PNEUMONIAE

SUBMISSION OF CERTIFIED COPY OF FOREIGN PRIORITY DOCUMENT

MAIL STOP AMENDMENT
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Applicants submit herewith a certified copy of the foreign priority document GB 9816336.3 as filed on 27 July 1998. No fee is believed due as a result of this submission. However, if a fee is due upon the filing of this document, please charge such fee to the undersigned's **Deposit Account No. 50-0206**.

Respectfully submitted,

Dated: September 30, 2005

By: [signature]
Laurence H. Posorske, Ph.D.
Registration No. 34,698

Christopher J. Nichols, Ph.D.
Registration No. 55,984

HUNTON & WILLIAMS LLP
Intellectual Property Department
1900 K Street, N.W., Suite 1200
Washington, DC 20006-1109
(202) 955-1500 (telephone)
(202) 778-2201 (facsimile)

LHP/CJN/cdh

Best Available Copy

The Patent Office

PATENTS • DESIGNS • COPYRIGHT • TRADE MARKS

CHARTER MARK

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

CERTIFIED COPY OF PRIORITY DOCUMENT

he undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the mptroller-General, hereby certify that annexed hereto is a true copy of the documents as inally filed in connection with the patent application identified therein.

I certify that the attached copy of the request for grant of a Patent (Form 1/77) bears an dment, effected by this office, following a request by the applicant and agreed to by the C troller-General.

a rdance with the Patents (Companies Re-registration) Rules 1982, if a company named in is tificate and any accompanying documents has re-registered under the Companies Act 8 th the same name as that with which it was registered immediately before re-registration e the substitution as, or inclusion as, the last part of the name of the words "public limited p " or their equivalents in Welsh, references to the name of the company in this certificate a ccompanying documents shall be treated as references to the name with which it is so eg ed.

I cor ce with the rules, the words "public limited company" may be replaced by p.l.c., plc, P . o C.

Re istr n under the Companies Act does not constitute a new legal entity but merely su s th mpany to certain additional company law rules.

Signed Andrew Gersey

Dated 14 September 2005

Patents Form 1/77

THE PATENT OFFICE
27 JUL 1998

Patents Act 1977
(Rule 16)

1/77

28JUL98 E378889-2 D00056
P01/7700 25.00 - 9816336.3

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1.	Your reference	PWC/P21122GB
2.	Patent application number *(The Patent Office will fill in this part)*	27 JUL 1998 **9816336.3**
3.	Full name, address and postcode of the or of each patent applicant *(underline all surnames)*	Microbial Technics Limited ~~Cortecs (UK) Limited~~ 20 Trumpington Street ~~The Old Blue School~~ Cambridge ~~Lower Square~~ ~~Isleworth~~ CB2 1QA. ~~Middlesex TW7 6RL~~ CF 20/8/99
	Patents ADP number *(if you know it)*	~~[illegible]~~ 06832471001.
	If the applicant is a corporate body, give the country/state of its incorporation	
4.	Title of the invention	PROTEINS
5.	Name of your agent *(if you have one)*	
	"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*	KILBURN & STRODE 20 RED LION STREET LONDON WC1R 4PJ
	Patents ADP number *(if you know it)*	125001
6.	If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or each of these earlier applications and *(if you know it)* the or each application number	Country / Priority application number *(if you know it)* / Date of filing *(day / month / year)*
7.	If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application	Number of earlier application number / Date of filing *(day / month / year)*
8.	Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:* *a) any applicant named in part 3 is not an inventor, or* *b) there is an inventor who is not named as an applicant, or* *c) any named applicant is a corporate body.* *See note (d))*	Yes

Patents Form

9. Enter number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form :

Description : 59

Claim(s) : 4

Abstract :

Drawing(s) :

10. If you are also filing any of the following, state how many against each item.

Priority documents :

Translations of priority documents :

Statement of inventorship and right to grant of a patent *(Patents Form 7/77)* :

Request for preliminary examination and search *(Patents Form 9/77)* :

Request for substantive examination *(Patents form 10/77)* :

Any other documents *(please specify)* :

11. I/We request the grant of a patent on the basis of this application.

Signature Kilburn & Strode

Kilburn & Strode

Date 27th July 1998

12. Name and daytime telephone number of person to contact in the United Kingdom

Mr. Paul W. Chapman
Tel: 0171-539 4200

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.*

b) *Write your answers in capital letters using black ink or you may type them.*

c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*

d) *If you have answered `Yes' Patents Form 7/77 will need to be filed.*

e) *Once you have filled in the form you must remember to sign and date it.*

f) *For details of the fee and ways to pay please contact the Patent Office.*

PROTEINS

The present invention relates to proteins derived from *Streptococcus pneumoniae*, nucleic acid molecules encoding such proteins, the use of the nucleic acid and/or proteins as antigens/immunogens and in detection/diagnosis, as well as methods for screening the proteins/nucleic acid sequences as potential anti-microbial targets.

Streptococcus pneumoniae, commonly referred to as the pneumococcus, is an important pathogenic organism. The continuing significance of *Streptoccocus pneumoniae* infections in relation to human disease in developing and developed countries has been authoritatively reviewed (Fiber, G.R., *Science*, **265**: 1385-1387 (1994)). That indicates that on a global scale this organism is believed to be the most common bacterial cause of acute respiratory infections, and is estimated to result in 1 million childhood deaths each year, mostly in developing countries (Stansfield, S.K., *Pediatr. Infect. Dis.*, **6**: 622 (1987)). In the USA it has been suggested (Breiman *et al*, *Arch. Intern. Med.*, **150**: 1401 (1990)) that the pneumococcus is still the most common cause of bacterial pneumonia, and that disease rates are particularly high in young children, in the elderly, and in patients with predisposing conditions such as asplenia, heart, lung and kidney disease, diabetes, alcoholism, or with immunosupressive disorders, especially AIDS. These groups are at higher risk of pneumococcal septicaemia and hence meningitis and therefore have a greater risk of dying from pneumococcal infection. The pneumococcus is also the leading cause of otitis media and sinusitis, which remain prevalent infections in children in developed countries, and which incur substantial costs.

The need for effective preventative strategies against pneumococcal infection is highlighted by the recent emergence of penicillin-resistant pneumococci. It has been reported that 6.6% of pneumoccal isolates in 13 US hospitals in 12 states were found to be resistant to penicillin and some isolates were also resistant to other antibiotics including third generation cyclosporins (Schappert, S.M., *Vital and Health Statistics of the Centres for Disease Control/National Centre for Health Statistics*, **214**:1 (1992)).

The rates of penicillin resistance can be higher (up to 20%) in some hospitals (Breiman *et al*, J. Am. Med. Assoc., **271**: 1831 (1994)). Since the development of penicillin resistance among pneumococci is both recent and sudden, coming after decades during which penicillin remained an effective treatment, these findings are regarded as alarming.

For the reasons given above, there are therefore compelling grounds for considering improvements in the means of preventing, controlling, diagnosing or treating pneumococcal diseases.

Various approaches have been taken in order to provide vaccines for the prevention of pneumococcal infections. Difficulties arise for instance in view of the variety of serotypes (at least 90) based on the structure of the polysaccharide capsule surrounding the organism. Vaccines against individual serotypes are not effective against other serotypes and this means that vaccines must include polysaccharide antigens from a whole range of serotypes in order to be effective in a majority of cases. An additional problem arises because it has been found that the capsular polysaccharides (each of which determines the serotype and is the major protective antigen) when purified and used as a vaccine do not reliably induce protective antibody responses in children under two years of age, the age group which suffers the highest incidence of invasive pneumococcal infection and meningitis.

A modification of the approach using capsule antigens relies on conjugating the polysaccharide to a protein in order to derive an enhanced immune response, particularly by giving the response T-cell dependent character. This approach has been used in the development of a vaccine against *Haemophilus influenzae*, for instance. There are, however, issues of cost concerning both the multi-polysaccharide vaccines and those based on conjugates.

A third approach is to look for other antigenic components which offer the potential to

be vaccine candidates. This is the basis of the present invention. Using a specially developed bacterial expression system, we have been able to identify a group of protein antigens from pneomococcus which are associated with the bacterial envelope or which are secreted.

Thus, in a first aspect the present invention provides a *Streptococcus pneumoniae* protein or polypeptide having a sequence selected from those shown in table 1.

In a second aspect, the present invention provides a *Streptococcus pneumoniae* protein or polypeptide having a sequence selected from those shown in table 2.

A protein or polypeptide of the present invention may be provided in substantially pure form. For example, it may be provided in a form which is substantially free of other proteins.

As discussed herein, the proteins and polypeptides of the invention are useful as antigenic material. Such material can be "antigenic" and/or "immunogenic". Generally, "antigenic" is taken to mean that the protein or polypeptide is capable of being used to raise antibodies or indeed is capable of inducing an antibody response in a subject. "Immunogenic" is taken to mean that the protein or polypeptide is capable of eliciting a protective immune response in a subject. Thus, in the latter case, the protein or polypeptide may be capable of not only generating an antibody response but, in addition, a non-antibody based immune response.

The skilled person will appreciate that homologues or derivatives of the proteins or polypeptides of the invention will also find use in the context of the present invention, ie as antigenic/immunogenic material. Thus, for instance proteins or polypeptides which include one or more additions, deletions, substitutions or the like are encompassed by the present invention. In addition, it may be possible to replace one amino acid with another of similar "type". For instance replacing one hydrophobic amino acid with another.

One can use a program such as the CLUSTAL program to compare amino acid sequences. This program compares amino acid sequences and finds the optimal alignment by inserting spaces in either sequence as appropriate. It is possible to calculate amino acid identity or similarity (identity plus conservation of amino acid type) for an optimal alignment. A program like BLASTx will align the longest stretch of similar sequences and assign a value to the fit. It is thus possible to obtain a comparison where several regions of similarity are found, each having a different score. Both types of identity analysis are contemplated in the present invention.

In the case of homologues and derivatives, the degree of identity with a protein or polypeptide as described herein is less important than that the homologue or derivative should retain the antigenicity or immunogenicity of the original protein or polypeptide. However, suitably, homologues or derivatives having at least 60% similarity (as discussed above) with the proteins or polypeptides described herein are provided. Preferably, homologues or derivatives having at least 70% similarity, more preferably at least 80% similarity are provided. Most preferably, homologues or derivatives having at least 90% or even 95% similarity are provided.

In an alternative approach, the homologues or derivatives could be fusion proteins, incorporating moieties which render purification easier, for example by effectively tagging the desired protein or polypeptide. It may be necessary to remove the "tag" or it may be the case that the fusion protein itself retains sufficient antigenicity to be useful.

In an additional aspect of the invention there are provided antigenic/immunogenic fragments of the proteins or polypeptides of the invention, or of homologues or derivatives thereof.

For fragments of the proteins or polypeptides described herein, or of homologues or derivatives thereof, the situation is slightly different. It is well known that is possible to screen an antigenic protein or polypeptide to identify epitopic regions, ie those regions

which are responsible for the protein or polypeptide's antigenicity or immunogenicity. Methods for carrying out such screening are well known in the art. Thus, the fragments of the present invention should include one or more such epitopic regions or be sufficiently similar to such regions to retain their antigenic/immunogenic properties. Thus, for fragments according to the present invention the degree of identity is perhaps irrelevant, since they may be 100% identical to a particular part of a protein or polypeptide, homologue or derivative as described herein. The key issue, once again, is that the fragment retains the antigenic/immunogenic properties.

Thus, what is important for homologues, derivatives and fragments is that they possess at least a degree of the antigenicity/immunogenicity of the protein or polypeptide from which they are derived.

Gene cloning techniques may be used to provide a protein of the invention in substantially pure form. These techniques are disclosed, for example, in J. Sambrook *et al Molecular Cloning* 2nd Edition, Cold Spring Harbor Laboratory Press (1989). Thus, in a third aspect, the present invention provides a nucleic acid molecule comprising or consisting of a sequence which is:

(i) any of the DNA sequences set out in Table 1 or their RNA equivalents;

(ii) a sequence which is complementary to any of the sequences of (i);

(iii) a sequence which codes for the same protein or polypeptide, as those sequences of (i) or (ii);

(iv) a sequence which has substantial identity with any of those of (i), (ii) and (iii);

(v) a sequence which codes for a homologue, derivative or fragment of a protein as defined in Table 1.

In a fourth aspect the present invention provides a nucleic acid molecule comprising or consisting of a sequence which is:

(i) any of the DNA sequences set out in Table 2 or their RNA equivalents;

(ii) a sequence which is complementary to any of the sequences of (i);

(iii) a sequence which codes for the same protein or polypeptide, as those sequences of (i) or (ii);

(iv) a sequence which has substantial identity with any of those of (i), (ii) and (iii); or

(v) a sequence which codes for a homologue, derivative or fragment of a protein as defined in Table 2.

The nucleic acid molecules of the invention may include a plurality of such sequences, and/or fragments. The skilled person will appreciate that the present invention can include novel variants of those particular novel nucleic acid molecules which are exemplified herein. Such variants are encompassed by the present invention. These may occur in nature, for example because of strain variation. For example, additions, substitutions and/or deletions are included. In addition, and particularly when utilising microbial expression systems, one may wish to engineer the nucleic acid sequence by making use of known preferred codon usage in the particular organism being used for expression. Thus, synthetic or non-naturally occurring variants are also included within the scope of the invention.

The term "RNA equivalent" when used above indicates that a given RNA molecule has a sequence which is complementary to that of a given DNA molecule (allowing for the fact that in RNA "U" replaces "T" in the genetic code).

When comparing nucleic acid sequences for the purposes of determining the degree of homology or identity one can use programs such as BESTFIT and GAP (both from the Wisconsin Genetics Computer Group (GCG) software package) BESTFIT, for example, compares two sequences and produces an optimal alignment of the most similar segments. GAP enables sequences to be aligned along their whole length and finds the optimal alignment by inserting spaces in either sequence as appropriate. Suitably, in the context of the present invention when discussing identity of nucleic acid sequences, the comparison is made by alignment of the sequences along their whole length.

Preferably, sequences which have substantial identity have at least 50% sequence identity, desirably at least 75% sequence identity and more desirably at least 90 or at least 95% sequence identity with said sequences. In some cases the sequence identity may be 99% or above.

Desirably, the term "substantial identity" indicates that said sequence has a greater degree of identity with any of the sequences described herein than with prior art nucleic acid sequences.

It should however be noted that where a nucleic acid sequence of the present invention codes for at least part of a novel gene product the present invention includes within its scope all possible sequence coding for the gene product or for a novel part thereof.

The nucleic acid molecule may be in isolated or recombinant form. It may be incorporated into a vector and the vector may be incorporated into a host. Such vectors and suitable hosts form yet further aspects of the present invention.

Therefore, for example, by using probes based upon the nucleic acid sequences provided herein, genes in *Streptococcus pneumoniae* can be identified. They can then be excised using restriction enzymes and cloned into a vector. The vector can be introduced into a suitable host for expression.

Nucleic acid molecules of the present invention may be obtained from *S.pneumoniae* by the use of appropriate probes complementary to part of the sequences of the nucleic acid molecules. Restriction enzymes or sonication techniques can be used to obtain appropriately sized fragments for probing.

Alternatively PCR techniques may be used to amplify a desired nucleic acid sequence. Thus the sequence data provided herein can be used to design two primers for use in PCR so that a desired sequence, including whole genes or fragments thereof, can be targeted and then amplified to a high degree.

Typically primers will be at least 15-25 nucleotides long.

As a further alternative chemical synthesis may be used. This may be automated. Relatively short sequences may be chemically synthesised and ligated together to provide a longer sequence.

There is another group of proteins from *S.pneumoniae* which have been identified using the bacterial expression system described herein. These are known proteins from *S.pneumoniae,* which have not previously been identified as antigenic proteins. The amino acid sequences of this group of proteins, together with DNA sequences coding for them are shown in Table 3. These proteins, or homologues, derivatives and/or fragments thereof also find use as antigens/immunogens. Thus, in another aspect the present invention provides the use of a protein or polypeptide having a sequence selected from those shown in Tables 1-3, or homologues, derivatives and/or fragments thereof, as an immunogen/antigen.

In yet a further aspect the present invention provides an immunogenic/antigenic composition comprising one or more proteins or polypeptides selected from those whose sequences are shown in Tables 1-3, or homologues or derivatives thereof, and/or fragments of any of these. In preferred embodiments, the immunogenic/antigenic composition is a vaccine or is for use in a diagnostic assay.

In the case of vaccines suitable additional excipients, diluents, adjuvants or the like may be included. Numerous examples of these are well known in the art.

It is also possible to utilise the nucleic acid sequences shown in Tables 1-3 in the preparation of so-called DNA vaccines. Thus, the invention also provides a vaccine composition comprising one or more nucleic acid sequences as defined herein. DNA vaccines are described in the art (see for instance, Donnelly *et al* , *Ann. Rev. Immunol.,* **15**:617-648 (1997)) and the skilled person can use such art described techniques to produce and use DNA vaccines according to the present invention.

As already discussed herein the proteins or polypeptides described herein, their homologues or derivatives, and/or fragments of any of these, can be used in methods of detecting/diagnosing *S.pneumoniae*. Such methods can be based on the detection of antibodies against such proteins which may be present in a subject. Therefore the present invention provides a method for the detection/diagnosis of *S.pneumoniae* which comprises the step of bringing into contact a sample to be tested with at least one protein, or homologue, derivative or fragment thereof, as described herein. Suitably, the sample is a biological sample, such as a tissue sample or a sample of blood or saliva obtained from a subject to be tested.

In an alternative approach, the proteins described herein, or homologues, derivatives and/or fragments thereof, can be used to raise antibodies, which in turn can be used to detect the antigens, and hence *S.pneumoniae*. Such antibodies form another aspect of

the invention. Antibodies within the scope of the present invention may be monoclonal or polyclonal.

Polyclonal antibodies can be raised by stimulating their production in a suitable animal host (e.g. a mouse, rat, guinea pig, rabbit, sheep, goat or monkey) when a protein as described herein, or a homologue, derivative or fragment thereof, is injected into the animal. If desired, an adjuvant may be administered together with the protein. Well-known adjuvants include Freund's adjuvant (complete and incomplete) and aluminium hydroxide. The antibodies can then be purified by virtue of their binding to a protein as described herein.

Monoclonal antibodies can be produced from hybridomas. These can be formed by fusing myeloma cells and spleen cells which produce the desired antibody in order to form an immortal cell line. Thus the well-known Kohler & Milstein technique (*Nature* **256** (1975)) or subsequent variations upon this technique can be used.

Techniques for producing monoclonal and polyclonal antibodies that bind to a particular polypeptide/protein are now well developed in the art. They are discussed in standard immunology textbooks, for example in Roitt *et al*, *Immunology* second edition (1989), Churchill Livingstone, London.

In addition to whole antibodies, the present invention includes derivatives thereof which are capable of binding to proteins etc as described herein. Thus the present invention includes antibody fragments and synthetic constructs. Examples of antibody fragments and synthetic constructs are given by Dougall *et al* in *Tibtech* **12** 372-379 (September 1994).

Antibody fragments include, for example, Fab, $F(ab')_2$ and Fv fragments. Fab fragments (These are discussed in Roitt *et al* [*supra*]). Fv fragments can be modified to produce a synthetic construct known as a single chain Fv (scFv) molecule. This includes a peptide

linker covalently joining V_h and V_l regions, which contributes to the stability of the molecule. Other synthetic constructs that can be used include CDR peptides. These are synthetic peptides comprising antigen-binding determinants. Peptide mimetics may also be used. These molecules are usually conformationally restricted organic rings that mimic the structure of a CDR loop and that include antigen-interactive side chains.

Synthetic constructs include chimaeric molecules. Thus, for example, humanised (or primatised) antibodies or derivatives thereof are within the scope of the present invention. An example of a humanised antibody is an antibody having human framework regions, but rodent hypervariable regions. Ways of producing chimaeric antibodies are discussed for example by Morrison *et al* in PNAS, **81,** 6851-6855 (1984) and by Takeda *et al* in Nature. **314,** 452-454 (1985).

Synthetic constructs also include molecules comprising an additional moiety that provides the molecule with some desirable property in addition to antigen binding. For example the moiety may be a label (e.g. a fluorescent or radioactive label). Alternatively, it may be a pharmaceutically active agent.

Antibodies, or derivatives thereof, find use in detection/diagnosis of *S.pneumoniae*. Thus, in another aspect the present invention provides a method for the detection/diagnosis of *S.pneumoniae* which comprises the step of bringing into contact a sample to be tested and antibodies capable of binding to one or more proteins described herein, or to homologues, derivatives and/or fragments thereof.

In addition, so-called "Affibodies" may be utilised. These are binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain (Nord *et al*.,). Thus, Small protein domains, capable of specific binding to different target proteins can be selected using combinatorial approaches.

It will also be clear that the nucleic acid sequences described herein may be used to detect/diagnose *S.pneumoniae*. Thus, in yet a further aspect, the present invention provides a method for the detection/diagnosis of *S.pneumoniae* which comprises the step of bringing into contact a sample to be tested with at least one nucleic acid sequence as described herein. Suitably, the sample is a biological sample, such as a tissue sample or a sample of blood or saliva obtained from a subject to be tested. Such samples may be pre-treated before being used in the methods of the invention. Thus, for example, a sample may be treated to extract DNA. Then, DNA probes based on the nucleic acid sequences described herein (ie usually fragments of such sequences) may be used to detect nucleic acid from *S.pneumoniae*.

In additional aspects, the present invention provides:

(a) a method of vaccinating a subject against *S.pneumoniae* which comprises the step of administering to a subject a protein or polypeptide of the invention, or a derivative, homologue or fragment thereof, or an immunogenic composition of the invention;

(b) a method of vaccinating a subject against *S.pneumoniae* which comprises the step of administering to a subject a nucleic acid molecule as defined herein;

(c) a method for the prophylaxis or treatment of *S.pneumoniae* infection which comprises the step of administering to a subject a protein or polypeptide of the invention, or a derivative, homologue or fragment thereof, or an immunogenic composition of the invention;

(d) a method for the prophylaxis or treatment of *S.pneumoniae* infection which comprises the step of administering to a subject a nucleic acid molecule as defined herein;

(e) a kit for use in detecting/diagnosing *S.pneumoniae* infection comprising one or more proteins or polypeptides of the invention, or homologues, derivatives or fragments thereof, or an antigenic composition of the invention; and

(f) a kit for use in detecting/diagnosing *S.pneumoniae* infection comprising one or more nucleic acid molecules as defined herein.

Given that we have identified a group of important proteins, such proteins are potential targets for anti-microbial therapy. It is necessary, however, to determine whether each individual protein is essential for the organism's viability. Thus, the present invention also provides a method of determining whether a protein or polypeptide as described herein represents a potential anti-microbial target which comprises antagonising, inhibiting or otherwise interfering with the function or expression of said protein and determining whether *S.pneumoniae* is still viable.

A suitable method for inactivating the protein is to effect selected gene knockouts, ie prevent expression of the protein and determine whether this results in a lethal change. Suitable methods for carrying out such gene knockouts are described in Li *et al*, *P.N.A.S.*, **94**:13251-13256 (1997) and Kolkman *et al*, **178**:3736-3741 (1996).

In a final aspect the present invention provides the use of an agent capable of antagonising, inhibiting or otherwise interfering with the function or expression of a protein or polypeptide of the invention in the manufacture of a medicament for use in the treatment or prophylaxis of *S.pneumoniae* infection.

As mentioned above, we have used a bacterial expression system as a means of identifying those proteins which are surface associated, secreted or exported and thus, would find use as antigens.

The information necessary for the secretion/export of proteins has been extensively studied in bacteria. In the majority of cases, protein export requires a signal peptide to be present at the N-terminus of the precursor protein so that it becomes directed to the translocation machinery on the cytoplasmic membrane. During or after translocation, the signal peptide is removed by a membrane associated signal peptidase. Ultimately the localization of the protein (i.e. whether it be secreted, an integral membrane protein or attached to the cell wall) is determined by sequences other than the leader peptide itself.

We are specifically interested in surface located or exported proteins as these are likely to be antigens for use in vaccines, as diagnostic reagents or as targets for therapy with novel chemical entities. We have therefore developed a screening vector-system in *Lactococcus lactis* that permits genes encoding exported proteins to be identified and isolated. We provide below a representative example showing how given novel surface associated proteins from *Streptococcus pneumoniae* have been identified and characterized. The screening vector incorporates the staphylococcal nuclease gene *nuc* lacking its own export signal as a secretion reporter. Staphylococcal nuclease is a naturally secreted heat-stable, monomeric enzyme which has been efficiently expressed and secreted in a range of Gram positive bacteria (Shortle, *Gene,* **22:**181-189 (1983); Kovacevic *et al.*, *J. Bacteriol.*, **162:**521-528 (1985); Miller *et al.*, *J. Bacteriol.*, **169:**3508-3514 (1987); Liebl *et al.*, *J. Bacteriol.,* **174:**1854-1861 (1992); Le Loir *et al.*, *J. Bacteriol.,* **176:**5135-5139 (1994); Poquet *et al.*, *J. Bacteriol.,* **180:**1904-1912 (1998)).

Recently, Poquet *et al.* ((1998), *supra*) have described a screening vector incorporating the *nuc* gene lacking its own signal leader as a reporter to identify exported proteins in Gram positive bacteria, and have applied it to *L. lactis*. This vector (pFUN) contains the pAMβ1 replicon which functions in a broad host range of Gram-positive bacteria in addition to the ColE1 replicon that promotes replication

in *Escherichia coli* and certain other Gram negative bacteria. Unique cloning sites present in the vector can be used to generate transcriptional and translational fusions between cloned genomic DNA fragments and the open reading frame of the truncated nuc gene devoid of its own signal secretion leader. The *nuc* gene makes an ideal reporter gene because the secretion of nuclease can readily be detected using a simple and sensitive plate test: Recombinant colonies secreting the nuclease develop a pink halo whereas control colonies remain white (Shortle, (1983), *supra;* Le Loir *et al.*, (1994), *supra*).

Thus, the invention will now be described with reference to the following representative example, which provides details of how the proteins, polypeptides and nucleic acid sequences described herein identified as antigenic targets.

We describe herein the construction of three reporter vectors and their use in *L. lactis* to identify and isolate genomic DNA fragments from *Streptococcus pneumoniae* encoding secreted or surface associated proteins.

EXAMPLE 1

(i) Construction of the pTREP1-nuc series of reporter vectors

(a) **Construction of expression plasmid pTREP1**

The pTREP1 plasmid is a high-copy number (40-80 per cell) theta-replicating gram positive plasmid, which is a derivative of the pTREX plasmid which is itself a derivative of the previously published pIL253 plasmid. pIL253 incorporates the broad Gram-positive host range replicon of pAMβ1 (Simon and Chopin, *Biochimie,* **70**:559-567 (1988)) and is non-mobilisable by the *L lactis* sex-factor. pIL253 also lacks the *tra* function which is necessary for transfer or efficient mobilisation by

conjugative parent plasmids exemplified by pIL501. The Enterococcal pAMβ1 replicon has previously been transferred to various species including *Streptococcus, Lactobacillus* and *Bacillus* species as well as *Clostridium acetobutylicum*, (Oultram and Klaenhammer, *FEMS Microbiological Letters,* **27**:129-134 (1985); Gibson *et al.*, {**FULL REF NEEDED**}1979; LeBlanc *et al.*, *Proceedings of the National Academy of Science USA,* **75**:3484-3487 (1978)) indicating the potential broad host range utility. The pTREP1 plasmid represents a constitutive transcription vector.

The pTREX vector was constructed as follows. An artificial DNA fragment containing a putative RNA stabilising sequence, a translation initiation region (TIR), a multiple cloning site for insertion of the target genes and a transcription terminator was created by annealing 2 complementary oligonucleotides and extending with Tfl DNA polymerase. The sense and anti-sense oligonucleotides contained the recognition sites for NheI and BamHI at their 5' ends respectively to facilitate cloning. This fragment was cloned between the XbaI and BamHI sites in pUC19NT7, a derivative of pUC19 which contains the T7 expression cassette from pLET1 (Wells *et al*, *J. Appl. Bacteriol.*, **74**:629-636 (1993)) cloned between the EcoRI and HindIII sites. The resulting construct was designated pUCLEX. The complete expression cassette of pUCLEX was then removed by cutting with HindIII and blunting followed by cutting with EcoRI before cloning into EcoRI and SacI (blunted) sites of pIL253 to generate the vector pTREX (Wells and Schofield, *In Current advances in metabolism, genetics and applications-NATO ASI Series,* **H 98**:37-62 (1996)). The putative RNA stabilising sequence and TIR are derived from the *Escherichia coli* T7 bacteriophage sequence and modified at one nucleotide position to enhance the complementarity of the Shine Dalgarno (SD) motif to the ribosomal 16s RNA of *Lactococcus lactis* (Schofield *et al*. pers. coms. University of Cambridge Dept. Pathology.).

A *Lactococcus lactis* MG1363 chromosomal DNA fragment exhibiting promoter activity which was subsequently designated P7 was cloned between the EcoRI and BglII sites present in the expression cassette, creating pTREX7. This active promoter region had been previously isolated using the promoter probe vector pSB292 (Waterfield *et al*, *Gene*, **165**:9-15 (1995)). The promoter fragment was amplified by PCR using the Vent DNA polymerase according to the manufacturer.

The pTREP1 vector was then constructed as follows. An artificial DNA fragment which included a transcription terminator, the forward pUC sequencing primer, a promoter multiple -cloning site region and a universal translation stop sequence was created by annealing two overlapping partially complementary synthetic oligonucleotides together and extending with sequenase according to manufacturers instructions. The sense and anti-sense ($pTREP_F$ and $pTREP_R$) oligonucleotides contained the recognition sites for EcoRV and BamHI at their 5' ends respectively to facilitate cloning into pTREX7. The transcription terminator was that of the *Bacillus penicillinase* gene, which has been shown to be effective in Lactococcus (Jos *et al.*, *Applied and Environmental Microbiology,* **50**:540-542 (1985)). This was considered necessary as expression of target genes in the pTREX vectors was observed to be leaky and is thought to be the result of cryptic promoter activity in the origin region (Schofield *et al.* pers. coms. University of Cambridge Dept. Pathology.). The forward pUC primer sequencing was included to enable direct sequencing of cloned DNA fragments. The translation stop sequence which encodes a stop codon in 3 different frames was included to prevent translational fusions between vector genes and cloned DNA fragments. The pTREX7 vector was first digested with EcoRI and blunted using the 5' - 3' polymerase activity of T4 DNA polymerase (NEB) according to manufacturer's instructions. The EcoRI digested and blunt ended pTREX7 vector was then digested with Bgl II thus removing the P7 promoter. The artificial DNA fragment derived from the annealed synthetic oligonucleotides was

then digested with EcoRV and Bam HI and cloned into the EcoRI(blunted)-Bgl II digested pTREX7 vector to generate pTREP. A *Lactococcus lactis* MG1363 chromosomal promoter designated P1 was then cloned between the EcoRI and BglII sites present in the pTREP expression cassette forming pTREP1. This promoter was also isolated using the promoter probe vector pSB292 and characterised by Waterfield *et al.*, (1995), *supra*. The P1 promoter fragment was originally amplified by PCR using vent DNA polymerase according to manufacturers instructions and cloned into the pTREX as an EcoRI-BglII DNA fragment. The EcoRI-BglII P1 promoter containing fragment was removed from pTREX1 by restriction enzyme digestion and used for cloning into pTREP (Schofield *et al.* pers. coms. University of Cambridge, Dept. Pathology.).

(b) <u>PCR amplification of the *S. aureus nuc* gene</u>.

The nucleotide sequence of the *S. aureus nuc* gene (EMBL database accession number V01281) was used to design synthetic oligonucleotide primers for PCR amplification. The primers were designed to amplify the mature form of the nuc gene designated nucA which is generated by proteolytic cleavage of the N-terminal 19 to 21 amino acids of the secreted propeptide designated Snase B (Shortle, (1983), *supra*). Three sense primers (nucS1, nucS2 and nucS3, Appendix 1) were designed, each one having a blunt-ended restriction endonuclease cleavage site for EcoRV or SmaI in a different reading frame with respect to the nuc gene. Additionally BglII and BamHI were incorporated at the 5' ends of the sense and anti-sense primers respectively to facilitate cloning into BamHI and BglII cut pTREP1. The sequences of all the primers are given in Appendix 1. Three nuc gene DNA fragments encoding the mature form of the nuclease gene (NucA) were amplified by PCR using each of the sense primers combined with the anti-sense primer described above. The nuc gene fragments were amplified by PCR using *S. aureus* genomic DNA template,

Vent DNA Polymerase (NEB) and the conditions recommended by the manufacturer. An initial denaturation step at 93 ^{O}C for 2 min was followed by 30 cycles of denaturation at 93 ^{O}C for 45 sec, annealing at 50 ^{O}C for 45 seconds, and extension at 73 ^{O}C for 1 minute and then a final 5 min extension step at 73 ^{O}C. The PCR amplified products were purified using a Wizard clean up column (Promega) to remove unincorporated nucleotides and primers.

(c) **Construction of the pTREP1-nuc vectors**

The purified nuc gene fragments described in section b were digested with Bgl II and BamHI using standard conditions and ligated to BamHI and BglII cut and dephosphorylated pTREP1 to generate the pTREP1-nuc1, pTREP1-nuc2 and pTREP1-nuc3 series of reporter vectors. General molecular biology techniques were carried out using the reagents and buffer supplied by the manufacture or using standard conditions(Sambrook and Maniatis, (1989), *supra*). In each of the pTREP1-nuc vectors the expression cassette comprises a transcription terminator, lactococcal promoter P1, unique cloning sites (BglII, EcoRV or SmaI) followed by the mature form of the nuc gene and a second transcription terminator. Note that the sequences required for translation and secretion of the nuc gene were deliberately excluded in this construction. Such elements can only be provided by appropriately digested foreign DNA fragments (representing the target bacterium) which can be cloned into the unique restriction sites present immediately upstream of the *nuc* gene.

In possessing a promoter, the pTREP1-nuc vectors differ from the pFUN vector described by Poquet *et al.* (1998), *supra*, which was used to identify *L. lactis* exported proteins by screening directly for Nuc activity directly in *L. lactis*. As the pFUN vector does not contain a promoter upstream of the *nuc* open reading frame the cloned genomic DNA fragment must also provide the signals for transcription in

addition to those elements required for translation initiation and secretion of Nuc. This limitation may prevent the isolation of genes that are distant from a promoter for example genes which are within polycistronic operons. Additionally there can be no guarantee that promoters derived from other species of bacteria will be recognised and functional in *L. lactis*. Certain promoters may be under stringent regulation in the natural host but not in *L. lactis*. In contrast, the presence of the P1 promoter in the pTREP1-nuc series of vectors ensures that promoterless DNA fragments (or DNA fragments containing promoter sequences not active in *L. lactis*) will still be transcribed.

(d) Screening for secreted proteins in *S. pneumoniae*

Genomic DNA isolated from *S. pneumoniae* was digested with the restriction enzyme Tru9I. This enzyme which recognises the sequence 5'- TTAA -3' was used because it cuts A/T rich genomes efficiently and can generate random genomic DNA fragments within the preferred size range (usually averaging 0.5 - 1.0 kb). This size range was preferred because there is an increased probability that the P1 promoter can be utilised to transcribe a novel gene sequence. However, the P1 promoter may not be necessary in all cases as it is possible that many Streptococcal promoters are recognised in *L. lactis*. DNA fragments of different size ranges were purified from partial Tru9I digests of *S. pneumoniae* genomic DNA. As the Tru 9I restriction enzyme generates staggered ends the DNA fragments had to be made blunt ended before ligation to the EcoRV or SmaI cut pTREP1-nuc vectors. This was achieved by the partial fill-in enzyme reaction using the 5'-3' polymerase activity of Klenow enzyme. Briefly Tru9I digested DNA was dissolved in a solution (usually between 10-20 μl in total) supplemented with T4 DNA ligase buffer (New England Biolabs; NEB) (1X) and 33 μM of each of the required dNTPs, in this case dATP and dTTP. Klenow enzyme was added (1 unit Klenow enzyme (NEB) per μg of

DNA) and the reaction incubated at 25°C for 15 minutes. The reaction was stopped by incubating the mix at 75°C for 20 minutes. EcoRV or SmaI digested pTREP-nuc plasmid DNA was then added (usually between 200-400 ng). The mix was then supplemented with 400 units of T4 DNA ligase (NEB) and T4 DNA ligase buffer (1X) and incubated overnight at 16°C. The ligation mix was precipitated directly in 100% Ethanol and 1/10 volume of 3M sodium acetate (pH 5.2) and used to transform *L. lactis* MG1363 (Gasson, 1983). Alternatively, the gene cloning site of the pTREP-nuc vectors also contains a BglII site which can be used to clone for example Sau3AI digested genomic DNA fragments.

L. lactis transformant colonies were grown on brain heart infusion agar and nuclease secreting (Nuc^+) clones were detected by a toluidine blue-DNA-agar overlay (0.05 M Tris pH 9.0, 10 g of agar per litre, 10 g of NaCl per liter, 0.1 mM CaCl2, 0.03% wt/vol. salmon sperm DNA and 90 mg of Toluidine blue O dye) essentially as described by Shortle, 1983, *supra* and Le Loir *et al.*, 1994, *supra*). The plates were then incubated at 37°C for up to 2 hours. Nuclease secreting clones develop an easily identifiable pink halo. Plasmid DNA was isolated from Nuc^+ recombinant *L. lactis* clones and DNA inserts were sequenced on one strand using the NucSeq sequencing primer described in Appendix 1, which sequences directly through the DNA insert.

Isolation of Genes Encoding Exported Proteins from *S. pneumoniae*

A large number of gene sequences putatively encoding exported proteins in *S. pneumoniae* have been identified using the nuclease screening system. These have now been further analysed to remove artefacts. The sequences identified using the screening system have been analysed using a number of parameters.

1. All putative surface proteins were analysed for leader/signal peptide sequences using the software programs Sequencher (Gene Codes Corporation) and DNA Strider (Marck, *Nucleic Acids Res.*, 16:1829-1836 (1988)). Bacterial signal peptide sequences share a common design. They are characterised by a short positively charged N-terminus (N region) immediately preceding a stretch of hydrophobic residues (central portion-h region) followed by a more polar C-terminal portion which contains the cleavage site (c-region). Computer software is available which allows hydropathy profiling of putative proteins and which can readily identify the very distinctive hydrophobic portion (h-region) typical of leader peptide sequences. In addition, the sequences were checked for the presence of or absence of a potential ribosomal binding site (Shine-Dalgarno motif) required for translation initiation of the putative nuc reporter fusion protein.

2. All putative surface protein sequences were also matched with all of the protein/DNA sequences using the publicly databases [OWL-proteins inclusive of SwissProt and GenBank translations]. This allows us to identify sequences similar to known genes or homologues of genes for which some function has been ascribed. Hence it has been possible to predict a function for some of the genes identified using the LEEP system and to unequivocally establish that the system can be used to identify and isolate gene sequences of surface associated proteins. We should also be able to confirm that these proteins are indeed surface related and not artifacts. The LEEP system has been used to identify novel gene targets for vaccine and therapy.

3. Some of the genes identified proteins did not possess a typical leader peptide sequence and did not show homology with any DNA/protein sequences in the database. Indeed these proteins may indicate the primary advantage of our screening method, i.e. the isolation of atypical surface-related proteins, which may have been missed in all previously described screening protocols or approaches based on sequence homology searches.

In all cases, only partial gene sequences were initially obtained. Full length genes were obtained in all cases by reference to the TIGR *S.pneumoniae* database (www@tigr.org). Thus, by matching the originally obtained partial sequences with the database, we were able to identify the full length gene sequences. In this way, as described herein, three groups of genes were clearly identified, ie a group of genes encoding previously unidentified *S.pneumoniae* proteins, a second group exhibiting some homology with known proteins from a variety of sources and a third group which encoded known *S.pneumoniae* proteins, which were, however, not known as antigens.

Appendix I - Oligonucleotide primers

nucS1

Bgl II Eco RV

5'- cgagatctgatatctcacaaacagataacggcgtaaatag -3'

nucS2

Bgl II Sma I

5'- gaagatcttcccgggatcacaaacagataacggcgtaaatag -3'

nucS3

Bgl II Eco RV

5'- cgagatctgatatccatcacaaacagataacggcgtaaatag -3'

nucR

Bam HI

5'- cgggatccttatggacctgaatcagcgttgtc -3'

NucSeq

5'- ggatgctttgtttcaggtgtatc -3'

pTREP$_F$

5'- catgatatcggtacctcaagctcatatcattgtccggcaatggtgtgggctttttttgttttagcggataa caatttcacac -3'

pTREP$_R$

5'- gcggatcccccgggcttaattaatgtttaaacactagtcgaagatctcgcgaattctcctgtgtgaaatt gttatccgcta -3'

pUC$_F$

5'- cgccagggttttcccagtcacgac -3'

V$_R$

5'- tcagggggcggagcctatg -3'

V$_1$

5'- tcgtatgttgtgtggaattgtg -3'

V_2

5'- tccggctcgtatgttgtgtggaattg -3'

TABLE 1

ID4 1200 bp

ATGAGAAATATGTGGGTTGTAATCAAGGAAACCTATCTTCGACATGTCGAGTCATGGAGTTTCTTCTTTATGGTGA
TTTCGCCGTTCCTCTTTTTAGGAATCTCTGTAGGAATTGGGCATCTCCAAGGTTCTTCTATGGCTAAAAATAATAAA
GTGGCAGTAGTGACAACAGTGCCATCTGTAGCAGAAGGACTGAAGAATGTAAATGGTGTTAACTTCGACTATAAA
GACGAAGCAAGTGCCAAAGAAGCAATTAAAGAAGAAAAATTAAAAGGTTATTTGACCATTGATCAAGAAGATAGT
GTTCTAAAGGCAGTTTATCATGGCGAAACATCGCTTGAAAATGGAATTAAATTTGAGGTTACAGGTACACTCAATG
AACTGCAAAATCAGCTTAATCGTTCAACTGCTTCCTTGTCTCAAGAGCAGGAAAAACGCTTAGCGCAGACAATTCA
ATTCACAGAAAAGATTGATGAAGCCAAGGAAAATAAAAAGTTTATTCAAACAATTGCAGCAGGTGCCTTAGGATT
CTTTCTTTATATGATTCTGATTACCTATGCGGGTGTAACAGCTCAGGAAGTTGCCAGTGAAAAAGGCACCAAAATT
ATGGAAGTCGTTTTTTCTAGCATAAGGGCAAGTCACTATTTCTATGCGCGGATGATGGCTCTGTTTCTAGTGATTTT
AACGCATATTGGGATCTATGTTGTAGGTGGTCTGGCTGCCGTTTTGCTCTTTAAAGATTTGCCATTCTTGGCTCAGT
CTGGTATTTTGGATCACTTGGGAGATGCTATCTCACTGAATACCTTGCTCTTTATTTTGATCAGTCTTTTCATGTAC
GTAGTCTTGGCAGCCTTCCTAGGATCTATGGTTTCTCGTCCTGAGGACTCAGGGAAAGCCTTGTCGCCTTTGATGA
TTTTGATTATGGGTGGTTTTTTTGGAGTGACAGCTCTAGGTGCAGCTGGTGACAATCTCCTCTTGAAGATTGGTTCT
TATATTCCCTTTATTTCGACCTTCTTTATGCCGTTTCGAACGATTAATGACTATGCGGGGGGAGCAGAAGCATGGA
TTTCACTTGCTATTACAGTGATTTTTGCGGTGGTAGCAACAGGATTTATCGGACGCATGTATGCTAGTCTCGTTCTT
CAAACGGATGATTTAGGGATTTGGAAAACCTTTAAACGTGCCTTATCTTATAAATAG

MRNMWVVIKETYLRHVESWSFFFMVISPFLFLGISVGIGHLQGSSMAKNNKVAVVTTVPSVAEGLKNVNGVNFDYKDE
ASAKEAIKEEKLKGYLTIDQEDSVLKAVYHGETSLENGIKFEVTGTLNELQNQLNRSTASLSQEQEKRLAQTIQFTEKIDE
AKENKKFIQTIAAGALGFFLYMILITYAGVTAQEVASEKGTKIMEVVFSSIRASHYFYARMMALFLVILTHIGIYVVGGL
AAVLLFKDLPFLAQSGILDHLGDAISLNTLLFILISLFMYVVLAAFLGSMVSRPEDSGKALSPLMILIMGGFFGVTALGAA
GDNLLLKIGSYIPFISTFFMPFRTINDYAGGAEAWISLAITVIFAVVATGFIGRMYASLVLQTDDLGIWKTFKRALSYKZ

ID5 1125 bp

CCTGGGAAAGTCTTGAAAATTATGATAGAATGGTGGAAGGAAAAATTCAGGAGAGTAGTAGTGACTCAAAATGTT
GAAAGTCTTCTCGTATCCATTGTAATCAGTGCATACAATGAAGAAAAATATCTGCCTGGTCTAATTGAAGACTTAA
AAAATCAAACCTATCCTAAAGAGGATATTGAAATTCTATTTATAAATGCTATGTCCACAGATGGACCACAGCTAT
CATTCAGCAATTTATAAAGGAAGATACAGAGTTTAACTCAATTAGATTGTATAACAATCCTAAGAAAAATCAAGC
TAGTGGTTTTAACCTGGGAGTTAAACATTCTGTAGGGGACCTTATTTTAAAAATTGATGCTCATTCAAAAGTTACT
GAGACTTTTGTAATGAACAATGTGGCTATTATTCAACAAGGTGAATTTGTCTGTGGGGGGCCTAGACCGACGATTG
TCGAAGGAAAAGGAAAATGGGCAGAGACCTTGCATCTTGTTGAGGAAAATATGTTTGGCAGTAGCATTGCCAATT
ATCGAAATAGTTCTGAGGATAGATATGTTTCTTCTATTTTTCATGGAATGTATAAACGAGAGGTTTTCCAGAAGGT
TGGTTTAGTAAATGAGCAACTTGGCCGAACTGAAGATAATGATATTCATTATAGAATTCGAGAATATGGTTATAAA
ATCCGCTATAGCCCAAGTATTCTATCTTATCAGTATATTCGACCAACATTCAAGAAAATGCTGCATCAAAAGTATT
CAAATGGTTTGTGGATTGGCTTGACAAGTCATGTTCAGTTTAAGTGTTTATCATTATTTCACTATGTTCCTTGTTTA
TTTGTTTTGAGTCTTGTGTTTAGTCTAGCATTGTTACCGATCACATTCGTATTCATAACTTTACTATTAGGTGCCTAT
TTTCTACTTTTGTCATTACTCACTTTGCTGACTTTATTAAAACATAAAAATGGATTTCTAATTGTGATGCCCTTTATT
TTATTTTCCATTCACTTTGCTTATGGCCTTGGGACGATTGTAGGTTTAATTAGAGGATTTAAATGGAAGAAGGAGT
ACAAGAGAACAATAATTTATTTGGATAAAATAAGCCAAATAAATCAAAATATGCTATAA

PGKVLKIMIEWWKEKFRRVVVTQNVESLLVSIVISAYNEEKYLPGLIEDLKNQTYPKEDIEILFINAMSTDGTTAIIQQFIK
EDTEFNSIRLYNNPKKNQASGFNLGVKHSVGDLILKIDAHSKVTETFVMNNVAIIQQGEFVCGGPRPTIVEGKGKWAET
LHLVEENMFGSSIANYRNSSEDRYVSSIFHGMYKREVFQKVGLVNEQLGRTEDNDIHYRIREYGYKIRYSPSILSYQYIRP
TFKKMLHQKYSNGLWIGLTSHVQFKCLSLFHYVPCLFVLSLVFSLALLPITFVFITLLLGAYFLLLSLLTLLTLLKHKNGF
LIVMPFILFSIHFAYGLGTIVGLIRGFKWKKEYKRTIIYLDKISQINQNMLZ

ID11 696 bp

ATCATGAAAGAACAAAATACGATAGAAATCGATGTATTTCAATTAGTTAAAAGCTTGTGGAAACGCAAGCTAATG
ATTTTAATAGTGGCACTTGTGACAGGTGCGGGGGCTTTTGCATATAGCACTTTTATTGTTAAGCCAGAATATACGA
GTACCACGCGAATTTACGTAGTGAATCGCAATCAAGGAGACAAGCCGGGGTTGACAAATCAGGATTTGCAGGCAG
GAACTTATCTGGTAAAAGACTACCGTGAGATTATCCTTTCGCAGGATGTTTTGGAGGAAGTTGTTTCTGATTTGAA
ACTAGATTTGACGCCAAAAGGTTTGGCTAATAAAATTAAAGTGACAGTACCAGTTGATACCCGTATTGTCTCTATT
TCAGTTAATGATCGAGTTCCTGAAGAGGCAAGCCGTATCGCTAACTCTTTGAGAGAAGTAGCTGCTCAAAAAATTA
TCAGTATTACTCGTGTTTCTGACGTGACAACACTGGAGGAGGCAAGGCCGGCGATATCCCCGTCTTCGCCAAATAT
TAAACGCAATACACTAATTGGTTTTTTGGCAGGGGTGATTGGAACTAGTGTTATAGTTCTTCATCTTGAACTTTTGG

ATACTCGTGTGAAACGTCCGGAAGATATCGAAAATACATTGCAGATGACACTTTTGGGAGTTGTGCCAAACTTGG
GTAAGTTGAAATAG

MMKEQNTIEIDVFQLVKSLWKRKLMILIVALVTGAGAFAYSTFIVKPEYTSTTRIYVVNRNQGDKPGLTNQDLQAGTYL
VKDYREIILSQDVLEEVVSDLKLDLTPKGLANKIKVTVPVDTRIVSISVNDRVPEEASRIANSLREVAAQKIISITRVSDVTT
LEEARPAISPSSPNIKRNTLIGFLAGVIGTSVIVLHLELLDTRVKRPEDIENTLQMTLLGVVPNLGKLKZ

ID19 555 bp

ATGGTAAAAGTAGCAGTTATATTAGCTCAGGGCTTTGAAGAAATTGAAGCCTTGACAGTTGTAGATGTCTTGCGTC
GAGCCAATATCACATGTGATATGGTTGGTTTTGAAGAGCAAGTAACGGGTTCGCATGCAATCCAAGTAAGAGCAG
ATCATGTCTTTGATGGAGATTTATCAGACTATGATATGATTGTTCTTCCTGGAGGTATGCCTGGTTCTGCACATTTA
CGTGATAATCAGACCTTGATTCAAGAATTGCAAAGCTTCGAGCAAGAAGGGAAGAAACTAGCAGCCATTTGTGCG
GCACCAATTGCCCTCAATCAAGCAGAGATATTGAAAAATAAGCGATACACTTGTTATGACGGCGTTCAAGAGCAA
ATCCTTGATGGTCACTACGTCAAGGAAACAGTAGTGGTAGATGGTCAGTTGACAACCAGTCGGGGTCCTTCAACA
GCCCTTGCCTTTGCCTACGAGTTGGTGGAGCAACTAGGAGGGGACGCAGAGAGTTTACGAACAGGAATGCTCTAT
CGAGATGTCTTTGGTAAAAATCAGTAA

MVKVAVILAQGFEEIEALTVVDVLRRANITCDMVGFEEQVTGSHAIQVRADHVFDGDLSDYDMIVLPGGMPGSAHLR
DNQTLIQELQSFEQEGKKLAAICAAPIALNQAEILKNKRYTCYDGVQEQILDGHYVKETVVVDGQLTTSRGPSTALAFA
YELVEQLGGDAESLRTGMLYRDVFGKNQZ

ID27 306 bp

GTGGTAGGGATGGTAGAACCAAACCTAGAAAGCCTTATAAAAGATCTTTACAATCATGCTCGACATGATTTGAGT
GAAGATTTAGTTGCTGCTCTCCTAGAGACTACTAAAAAACTGCCTACTACAAATGAGCAATTGCAGGCAGTTCGTC
TCTCAGGCCTGGTCAATCGTGAATTGCTCCTAAATCCCAAACATCCAGCACCTGAGTTGCTCAACTTGGCTCGCTT
TGTCAAAAGAGAAGAAGCCAAGTACAGAGGAACTGCGACTTCTGCGCTTATGTATGAGGAACTCTTTAAAATGCT
TTGA

MVGMVEPNLESLIKDLYNHARHDLSEDLVAALLETTKKLPTTNEQLQAVRLSGLVNRELLLNPKHPAPELLNLARFVK
REEAKYRGTATSALMYEELFKMLZ

ID29 945 bp

TTGTTCTTAAAAAAGGAAAGAGAGGTAATCAGCATGCGTAAATGGACAAAAGGATTTCTCATCTTTGGTGTGGTG
ACTACCGTTATCGGCTTTATCCTGCTTTTTGTAGGTATCCAATCTGACGGGAtTAAGAGCCTACTTTCCATGTCCAA
AGAACCTGTCTATGATAGCCGTACGGAAAAGCTAACCTTTGGCAAGGAAGTCGAAAACCTAGAAATTACTCTCCA
CCAACACACGCTCACCATCACAGACTCTTTCGATGATCAAATCCACATTTCTTACCATCCATCTCTTTCTGCTCACC
ATGATCTTATCACCAATCAGAACGATAGAACTCTGAGTCTCACTGATAAGAAACTGTCTGAAACTCCGTTTCTCTC
TTCTGGAATTGGTGGGATTCTTCATATCGCAAGTAGCTACTCTAGTCGTTTTGAAGAAGTTATTCTCCGACTACCA
AAAGGGAGAACTCTAAAAGGGATCAACATCTCAGCCAATCGCGGACAAACCACCATCATAAATGCTAGCCTTGAA
AATGCGACCCTCAATACAAACAGCTATATCCTCCGAATTGAAGGAAGTCGTATCAAAAACAGTAAACTCACAACG
CCCAATATCGTTAATATCTTTGATACAGTTCTTACAGATAGTCAGCTAGAGTCAACAGAGAATCACTTCCACGCTG
AAAATATCCAAGTCCATGGCAAGGTTGAACTGACTGCCAAAGATTATCTCAGAATCATCCTAGACCAGAAAGAAA
GCCAACGAATTAACTGGGACATCTCAAGCAACTATGGTTCTATCTTCCAATTCACAAGAGAAAAGCCTGAATCAA
GAGGTACGGAATTAAGCAACCCTTACAAAACTGAAAAAACCGATGTCAAGGATCAACTCATTGCGAGATCTGATG
ATAATATTGATCTAATATCCACACCAAGCAGACGTTGA

MFLKKEREVISMRKWTKGFLIFGVVTTVIGFILLFVGIQSDGIKSLLSMSKEPVYDSRTEKLTFGKEVENLEITLHQHTLTI
TDSFDDQIHISYHPSLSAHHDLITNQNDRTLSLTDKKLSETPFLSSGIGGILHIASSYSSRFEEVILRLPKGRTLKGINISANR
GQTTIINASLENATLNTNSYILRIEGSRIKNSKLTTPNIVNIFDTVLTDSQLESTENHFHAENIQVHGKVELTAKDYLRIILD
QKESQRINWDISSNYGSIFQFTREKPESRGTELSNPYKTEKTDVKDQLIARSDDNIDLISTPSRRZ

ID30 879 bp

ATGAAACAAGAATGGTTTGAAAGTAATGATTTTGTAAAAACAACAAGCAAGAACAAGCCTGAAGAGCAAGCTCAA
GAGGTTGCAGACAAGGCTGAAGAAACGATAGCCGATCTCGATACACCAATTGAAAAAAATACTCAGTTAGAGGAG
GAAGTCCCTCAAGCTGAAGTCGAATTGGAAAGCCAGCAAGAAGAGAAAATTGAAGCTCCTGAAGACAGTGAAGC
GAGAACAGAAATAGAAGAAAAGAAGGCATCTAATTCTACTGAAGAAGAGCCAGACCTTTCTAAAGAAACAGAAA
AAGTCACTATAGCTGAAGAGAGCCAAGAAGCTCTTCCTCAGCAAAAAGCAACCACGAAAGAGCCACTTCTTATCA
GTAAATCTTTAGAAAGTCCTTATATCCCCGACCAAGCTCCAAAATCTAGGGATAAATGGAAAGAGCAAGTGCTTG
ATTTTTGGTCTTGGCTAGTGGAAGCGATCAAATCTCCTACAAGTAAGTTGGAAACAAGTATCACACACAGTTACAC
AGCCTTTCTCTTGCTCATTCTGTTTTCTGCATCTTCCTTTTTCTTTAGTATCTATCACATCAAACATGCTTACTATGG

ACATATAGCAAGCATTAACAGTCGCTTCCCTGAGCAGCTAGCTCCTTTAACTCTTTTTTCTATCATCTCTATCCTAG
TAGCGACAACACTCTTCTTCTTTTCATTCCTCTTGGGTAGTTTCGTTGTGAGACGATTTATCCACCAGGAAAAGGA
CTGGACGCTAGACAAGGTTCTCCAACAATATAGTCAACTCTTGGCAATTCCAATCTCCTCACTGCTATTGCTAGTT
TCTTTGCTTTCTTTGATAGCCTACGATTTACAGCCCTCTTGTGTGTGA

MKQEWFESNDFVKTTSKNKPEEQAQEVADKAEETIADLDTPIEKNTQLEEEVPQAEVELESQQEEKIEAPEDSEARTEIE
EKKASNSTEEEPDLSKETEKVTIAEESQEALPQQKATTKEPLLISKSLESPYIPDQAPKSRDKWKEQVLDFWSWLVEAIKS
PTSKLETSITHSYTAFLLLILFSASSFFFSIYHIKHAYYGHIASINSRFPEQLAPLTLFSIISILVATTLFFFSFLLGSFVVRRFIH
QEKDWTLDKVLQQYSQLLAIPISSLLLLVSLLSLIAYDLQPSCVZ

ID105 990 bp

ATGCAACTCGCTTCTTCGGTCTACTCATTGTTCGTCTGGTACAATTTGTTCTTAAAAAAGGAAAGAGAGGTAATCA
GCATGCGTAAATGGACAAAAGGATTTCTCATCTTTGGTGTGGTGACTACCGTTATCGGCTTTATCCTGCTTTTTGTA
GGTATCCAATCTGACGGGATTAAGAGCCTACTTTCCATGTCCAAAGAACCTGTCTATGATAGCCGTACGGAAAAG
CTAACCTTTGGCAAGGAAGTCGAAAACCTAGAAATTACTCTCCACCAACACACGCTCACCATCACAGACTCTTTCG
ATGATCAAATCCACATTTCTTACCATCCATCTCTTTCTGCTCACCATGATCTTATCACCAATCAGAACGATAGAAC
TCTGAGTCTCACTGATAAGAAACTGTCTGAAACTCCGTTTCTCTCTTCTGGAATTGGTGGGATTCTTCATATCGCAA
GTAGCTACTCTAGTCGTTTTGAAGAAGTTATTCTCCGACTACCAAAAGGGAGAACTCTAAAAGGGATCAACATCTC
AGCCAATCGCGGACAAACCACCATCATAAATGCTAGCCTTGAAAATGCGACCCTCAATACAAACAGCTATATCCT
CCGAATTGAAGGAAGTCGTATCAAAAACAGTAAACTCACAACGCCCAATATCGTTAATATCTTTGATACAGTTCTT
ACAGATAGTCAGCTAGAGTCAACAGAGAATCACTTCCACGCTGAAAATATCCAAGTCCATGGCAAGGTTGAACTG
ACTGCCAAAGATTATCTCAGAATCATCCTAGACCAGAAAGAAAGCCAACGAATTAACTGGGACATCTCAAGCAAC
TATGGTTCTATCTTCCAATTCACAAGAGAAAAGCCTGAATCAAGAGGTACGGAATTAAGCAACCCTTACAAAACT
GAAAAAACCGATGTCAAGGATCAACTCATTGCGAGATCTGATGATAATATTGATCTAATATCCACACCAAGCAGA
CGTTGA

MQLASSVYSLFVWYNLFLKKEREVISMRKWTKGFLIFGVVTTVIGFILLFVGIQSDGIKSLLSMSKEPVYDSRTEKLTFGK
EVENLEITLHQHTLTITDSFDDQIHISYHPSLSAHHDLITNQNDRTLSLTDKKLSETPFLSSGIGGILHIASSYSSRFEEVILR
LPKGRTLKGINISANRGQTTIINASLENATLNTNSYILRIEGSRIKNSKLTTPNIVNIFDTVLTDSQLESTENHFHAENIQVH
GKVELTAKDYLRIILDQKESQRINWDISSNYGSIFQFTREKPESRGTELSNPYKTEKTDVKDQLIARSDDNIDLISTPSRRZ

ID107 –78bp

ATGATATGTAAAATGAAGCAGGGAGGGAGCAGGGCGTGCTGGGGATGGAGAGTGGGGGAGGGACGCTGCTATTTT
AATC

MICKMKQGGSRACWGWRVGEGRCYFN

ID109 714 bp

CGATAAAGAGGCCTTGAGTAATCTCAATTTGCAGATTGAAAATGGAGAGATTATGGGCTTGATTGGTCATAATGG
GGCTGGAAAATCGACCACTATAAAATCCCTAGTCAGTATCATTTCACCCAGCAGTGGTCGTATTTTGGTAGACGGT
CAGGAGTTATCGGAAAATCGCTTGGCTATTAAACGAAAGATTGGCTACGTAGCAGACTCGCCTGACTTATTTTTAC
GCTTAACGGCCAATGAATTTTGGGAATTGATCGCCTCATCCTATGATCTGAGTAGATCTGACTTGGAGGCTAGTCT
AGCTAGGCTATTGAACGTTTTTGATTTTGCTGAAAATCGCTATCAGGTTATTGAAACTCTTTCTCACGGAATGCGTC
AGAAAGTCTTTGTCATCGGAGCACTCTTGTCTGATCCCGATATTTGGGTTTTGGACGAACCCTTGACTGGTTTGGA
TCCCCAGGCTGCCTTTGATTTGAAACAGATGATGAAGGAACATGCACAAAAAGGGAAGACAGTCTTGTTTTCAAC
TCATGTCCTAGAGGTGGCAGAGCAAGTCTGTGATCGGATTGCCATTTTGAAAAAGGGGCATTTGATTTATTGTGGT
AAGGTAGAGGACTTGAGGAAAGACCACCCAGACCAGTCTTTGGAAAGTATCTACCTTAGTCTTGCTGGTAGAAAA
GAGGAGGTTGCGGATGCGTCTCAAGGTCATTAA

DKEALSNLNLQIENGEIMGLIGHNGAGKSTTIKSLVSIISPSSGRILVDGQELSENRLAIKRKIGYVADSPDLFLRLTANEF
WELIASSYDLSRSDLEASLARLLNVFDFAENRYQVIETLSHGMRQKVFVIGALLSDPDIWVLDEPLTGLDPQAAFDLKQ
MMKEHAQKGKTVLFSTHVLEVAEQVCDRIAILKKGHLIYCGKVEDLRKDHPDQSLESIYLSLAGRKEEVADASQGHZ

ID112 360 bp

ATGGCTTTGTTTTCAGAGAGAGGAGCAGTACGGAAGACACCAATGGCAAGTCCAATAATGAGACCTATGATGGTT
CCGACGATAGAGATTAAAAGAGTGATACCAGCACCACGCAAGAGTTGTTGCCAGTTTTCAGAAAGAATTTTAGCA
ACTTGGCTAAAGAAACTACTGCTAGTCTCTTCAGTTGTTGTAGCTTCGGCAGGTTGTTCCTTGATCATACGATCCAT

CAAGGCAACTTGGTCATCTTTTGAAATGGTTTCAATGCTGGCATTGATTTGGCTAATACGATTGTCATTTTTACGAA
GCCCGATAGCGATAGCTGTATCTTCTTCCCCAGTTTTGAAACCAGGTTCTACTTGA

MALFSERGAVRKTPMASPIMRPMMVPTIEIKRVIPAPRKSCCQFSERILATWLKKLLLVSSVVVASAGCSLIIRSIKATWSS
FEMVSMLALIWLIRLSFLRSPIAIAVSSSPVLKPGSTZ

TABLE 2

ID2 840 bp

ATGGGAATTGCTCTAGAAAATGTGAATTTTACATATCAAGAAGGTACTCCCTTAGCTTCAGCAGCTTTGTCGGATG
TTTCTTTGACGATTGAAGATGGCTCTTATACAGCTTTAATTGGGCACACAGGTAGTGGTAAATCAACTATTTTACA
ACTCTTAAATGGTTTATTGGTGCCAAGTCAAGGGAGTGTGAGGGTTTTTGATACCTTAATCACCTCGACTTCTAAA
AATAAAGATATTCGTCAAATTAGAAAACAGGTTGGCTTGGTATTTCAGTTTGCTGAAAATCAGATTTTTGAAGAAA
CGGTTTTGAAGGACGTTGCTTTTGGACCGCAAAATTTTGGAGTTTCTGAAGAAGATGCTGTGAAGACTGCGCGTGA
GAAACTGGCTCTGGTTGGAATTGATGAATCACTTTTTGATCGTAGTCCGTTTGAGCTGTCAGGGGGACAAATGAGA
CGTGTTGCCATTGCAGGCATACTTGCCATGGAGCCAGCTATATTAGTCTTAGATGAGCCAACAGCTGGTCTAGATC
CTCTAGGGAGAAAAGAGTTGATGACCCTGTTCAAAAAACTCCACCAGTCAGGGATGACCATCGTCTTGGTAACGC
ATTTGATGGATGATGTTGCTGAATATGCGAATCAAGTCTATGTAATGGAAAAGGGACGTTTAGTAAAGGGGGGCA
AACCAAGTGATGTCTTTCAAGACGTTGTTTTTATGGAAGAAGTTCAGTTGGGAGTACCTAAAATTACGGCCTTTTG
TAAACGATTGGCTGATAGAGGCGTGTCATTTAAACGATTACCGATTAAGATAGAGGAGTTCAAGGAGTCGCTAAA
TGGATAG

MGIALENVNFTYQEGTPLASAALSDVSLTIEDGSYTALIGHTGSGKSTILQLLNGLLVPSQGSVRVFDTLITSTSKNKDIR
QIRKQVGLVFQFAENQIFEETVLKDVAFGPQNFGVSEEDAVKTAREKLALVGIDESLFDRSPFELSGGQMRRVAIAGILA
MEPAILVLDEPTAGLDPLGRKELMTLFKKLHQSGMTIVLVTHLMDDVAEYANQVYVMEKGRLVKGGKPSDVFQDVV
FMEEVQLGVPKITAFCKRLADRGVSFKRLPIKIEEFKESLNGZ

ID 3 6360 bp

TACCCGGTAGTCTTAGCAGACACATCTAGCTCTGAAGATGCTTTAAACATCTCTGATAAAGAAAAAGTAGCAGAA
AATAAAGAGAAACATGAAAATATCCATAGTGCTATGGAAACTTCACAGGATTTTAAAGAGAAGAAAACAGCAGTC
ATTAAGGAAAAAGAAGTTGTTAGTAAAAATCCTGTGATAGACAATAACACTAGCAATGAAGAAGCAAAAATCAAA
GAAGAAAATTCCAATAAATCCCAAGGAGATTATACGGACTCATTTGTGAATAAAAACACAGAAAATCCCAAAAAA
GAAGATAAAGTTGTCTATATTGCTGAATTTAAAGATAAAGAATCTGGAGAAAAAGCAATCAAGGAACTATCCAGT
CTTAAGAATACAAAAGTTTTATATACTTATGATAGAATTTTTAACGGTAGTGCCATAGAAACAACTCCAGATAACT
TGGACAAAATTAAACAAATAGAAGGTATTTCATCGGTTGAAAGGGCACAAAAAGTCCAACCCATGATGAATCATG
CCAGAAAGGAAATTGGAGTTGAGGAAGCTATTGATTACCTAAAGTCTATCAATGCTCCGTTTGGGAAAAATTTTGA
TGGTAGAGGTATGGTCATTTCAAATATCGATACTGGAACAGATTATAGACATAAGGCTATGAGAATCGATGATGA
TGCCAAAGCCTCAATGAGATTTAAAAAAGAAGACTTAAAAGGCACTGATAAAAATTATTGGTTGAGTGATAAAAT
CCCTCATGCGTTCAATTATTATAATGGTGGCAAAATCACTGTAGAAAAATATGATGATGGAAGGGATTATTTTGAC
CCACATGGGATGCATATTGCAGGGATTCTTGCTGGAAATGATACTGAACAAGACATCAAAAACTTTAACGGCATA
GATGGAATTGCACCTAATGCACAAATTTTCTCTTACAAAATGTATTCTGACGCAGGATCTGGGTTTGCGGGTGATG
AAACAATGTTTCATGCTATTGAAGATTCTATCAAACACAACGTTGATGTTGTTTCGGTATCATCTGGTTTTACAGG
AACAGGTCTTGTAGGTGAGAAATATTGGCAAGCTATTCGGGCATTAAGAAAAGCAGGCATTCCAATGGTTGTCGC
TACGGGTAACTATGCGACTTCTGCTTCAAGTTCTTCATGGGATTTAGTAGCAAATAATCATCTGAAAATGACCGAC
ACTGGAAATGTAACACGAACTGCAGCACATGAAGATGCGATAGCGGTCGCTTCTGCTAAAAATCAAACAGTTGAG
TTTGATAAAGTTAACATAGGTGGAGAAAGTTTTAAATACAGAAATATAGGGGCCTTTTTCGATAAGAGTAAAATC
ACAACAAATGAAGATGGAACAAAAGCTCCTAGTAAATTAAAATTTGTATATATAGGCAAGGGGCAAGACCAAGAT
TTGATAGGTTTGGATCTTAGGGGCAAAATTGCAGTAATGGATAGAATTTATACAAAGGATTTAAAAAATGCTTTTA
AAAAAGCTATGGATAAGGGTGCACGCGCCATTATGGTTGTAAATACTGTAAATTACTACAATAGAGATAATTGGA
CAGAGCTTCCAGCTATGGGATATGAAGCGGATGAAGGTACTAAAAGTCAAGTGTTTTCAATTTCAGGAGATGATG
GTGTAAAGCTATGGAACATGATTAATCCTGATAAAAAAACTGAAGTCAAAAGAAATAATAAAGAAGATTTTAAAG
ATAAATTGGAGCAATACTATCCAATTGATATGGAAAGTTTTAATTCCAACAAACCGAATGTAGGTGACGAAAAAG
AGATTGACTTTAAGTTTGCACCTGACACAGACAAAGAACTCTATAAAGAAGATATCATCGTTCCAGCAGGATCTA
CATCTTGGGGGCCAAGAATAGATTTACTTTTAAAACCCGATGTTTCAGCACCTGGTAAAAATATTAAATCCACGCT
TAATGTTATTAATGGCAAATCAACTTATGGCTATATGTCAGGAACTAGTATGGCGACTCCAATCGTGGCAGCTTCT
ACTGTTTTGATTAGACCGAAATTAAAGGAAATGCTTGAAAGACCTGTATTGAAAAATCTTAAGGGAGATGACAAA
ATAGATCTTACAAGTCTTACAAAAATTGCCCTACAAAATACTGCGCGACCTATGATGGATGCAACTTCTTGGAAAG
AAAAAAGTCAATACTTTGCATCACCTAGACAACAGGGAGCAGGCCTAATTAATGTGGCCAATGCTTTGAGAAATG
AAGTTGTAGCAACTTTCAAAAACACTGATTCTAAAGGTTTGGTAAACTCATATGGTTCCATTTCTCTTAAAGAAAT
AAAAGGTGATAAAAAATACTTTACAATCAAGCTTCACAATACATCAAACAGACCTTTGACTTTTAAAGTTTCAGCA
TCAGCGATAACTACAGATTCTCTAACTGACAGATTAAAACTTGATGAAACATATAAAGATGAAAAATCTCCAGAT
GGTAAGCAAATTGTTCCAGAAATTCACCCAGAAAAAGTCAAAGGAGCAAATATCACATTTGAGCATGATACTTTC
ACTATAGGCGCAAATTCTAGCTTTGATTTGAATGCGGTTATAAATGTTGGAGAGGCCAAAAACAAAAATAAATTTG
TAGAATCATTTATTCATTTTGAGTCAGTGGAAGCGATGGAAGCTCTAAACTCCAGCGGGAAGAAAATAAACTTCC
AACCTTCTTTGTCGATGCCTCTAATGGGATTTGCTGGGAATTGGAACCACGAACCAATCCTTGATAAATGGGCTTG
GGAAGAAGGGTCAAGATCAAAAACACTGGGAGGTTATGATGATGATGGTAAACCGAAAATTCCAGGAACCTTAAA
TAAGGGAATTGGTGGAGAACATGGTATAGATAAATTTAATCCAGCAGGAGTTATACAAAATAGAAAAGATAAAAA
TACAACATCCCTGGATCAAAATCCAGAATTATTTGCTTTCAATAACGAAGGGATCAACGCTCCATCATCAAGTGGT

TCTAAGATTGCTAACATTTATCCTTTAGATTCAAATGGAAATCCTCAAGATGCTCAACTTGAAAGAGGATTAACAC
CTTCTCCACTTGTATTAAGAAGTGCAGAAGAAGGATTGATTTCAATAGTAAATACAAATAAAGAGGGAGAAAATC
AAAGAGACTTAAAAGTCATTTCGAGAGAACACTTTATTAGAGGAATTTTAAATTCTAAAAGCAATGATGCAAAGG
GAATCAAATCATCTAAACTAAAAGTTTGGGGTGACTTGAAGTGGGATGGACTCATCTATAATCCTAGAGGTAGAG
AAGAAAATGCACCAGAAAGTAAGGATAATCAAGATCCTGCTACTAAGATAAGAGGTCAATTTGAACCGATTGCGG
AAGGTCAATATTTCTATAAATTTAAATATAGATTAACTAAAGATTACCCATGGCAGGTTTCCTATATTCCTGTAAA
AATTGATAACACCGCCCCTAAGATTGTTTCGGTTGATTTTTCAAATCCTGAAAAAATTAAGTTGATTACAAAGGAT
ACTTATCATAAGGTAAAAGATCAGTATAAGAATGAAACGCTATTTGCGAGAGATCAAAAAGAACATCCTGAAAAA
TTTGACGAGATTGCGAACGAAGTTTGGTATGCTGGCGCCGCTCTTGTTAATGAAGATGGAGAGGTTGAAAAAAAT
CTTGAAGTAACTTACGCAGGTGAGGGTCAAGGAAGAAATAGAAAACTTGATAAAGACGGAAATACCATTTATGAA
ATTAAAGGTGCGGGAGATTTAAGGGGAAAAATCATTGAAGTCATTGCATTAGATGGTTCTAGCAATTTCACAAAG
ATTCATAGAATTAAATTTGCTAATCAGGCTGATGAAAAGGGGATGATTTCCTATTATCTAGTAGATCCTGATCAAG
ATTCATCTAAATATCAAAAGCTTGGCGAGATTGCAGAATCTAAATTTAAAAATTTAGGAAATGGAAAAGAGGGTA
GTCTAAAAAAAGATACAACTGGGGTAGAACATCATCATCAAGAAAATGAAGAGTCTATTAAAGAAAAATCTAGTT
TTACTATTGATAGAAATATTTCAACAATTAGAGACTTTGAAAATAAAGACTTAAAGAAACTCATTAAAAAGAAATT
TAGAGAAGTTGATGATTTTACAAGTGAAACTGGTAAGAGAATGGAGGAATACGATTATAAATACGATGATAAAGG
AAATATAATAGCCTACGATGATGGGACTGATCTAGAATATGAAACTGAGAAACTTGACGAAATCAAATCAAAAAT
TTATGGTGTTCTAAGTCCGTCTAAAGATGGACACTTTGAAATTCTTGGAAAGATAAGTAATGTTTCTAAAAATGCC
AAGGTATATTATGGGAATAACTATAAATCTATAGAAATCAAAGCGACCAAGTATGATTTCCACTCAAAAACGATG
ACATTTGATCTATACGCTAATATTAATGATATTGTGGATGGATTAGCTTTTGCAGGAGATATGAGATTATTTGTTAA
AGATAATGATCAGAAAAAAGCTGAAATTAAAATTAGAATGCCTGAAAAAATTAAGGAAACTAAATCAGAATATCC
CTATGTATCAAGTTATGGGAATGTCATAGAATTAGGGGAAGGAGATCTTTCAAAAAACAAACCAGACAATTTAAC
TAAAATGGAATCTGGTAAAATCTATTCTGATTCAGAAAAACAACAATATCTGTTAAAGGATAATATCATTCTAAGA
AAAGGCTATGCACTAAAAGTGACTACCTATAATCCTGGAAAAACGGATATGTTAGAAGGAAATGGAGTCTATAGC
AAGGAAGATATAGCAAAAATACAAAAGGCCAATCCTAATCTAAGAGCCCTTTCAGAAACAACAATTTATGCTGAT
AGTAGAAATGTTGAAGATGGAAGAAGTACCCAATCTGTATTAATGTCGGCTTTGGACGGCTTTAACATTATAAGGT
ATCAAGTGTTTACATTTAAAATGAACGATAAAGGGGAAGCTATCGATAAAGACGGAAATCTTGTGACAGATTCTT
CTAAACTTGTATTATTTGGTAAGGATGATAAAGAATACACTGGAGAGGATAAGTTCAATGTAGAAGCTATAAAAG
AAGATGGCTCCATGTTATTTATTGATACCAAACCAGTAAACCTTTCAATGGATAAGAACTACTTTAATCCATCTAA
ATCTAATAAAATTTATGTACGAAATCCAGAATTTTATTTAAGAGGTAAGATTTCTGATAAGGGTGGTTTTAACTGG
GAATTGAGAGTTAATGAATCGGTTGTAGATAATTATTTAATCTACGGAGATTTACACATTGATAACACTAGAGATT
TTAATATTAAGCTGAATGTTAAAGACGGTGACATCATGGACTGGGGAATGAAAGACTATAAAGCAAACGGATTTC
CAGATAAGGTAACAGATATGGATGGAAATGTTTATCTTCAAACTGGCTATAGCGATTTGAATGCTAAAGCAGTTGG
AGTCCACTATCAGTTTTTATATGATAATGTTAAACCCGAAGTAAACATTGATCCTAAGGGAAATACTAGTATCGAA
TATGCTGATGGAAAATCTGTAGTCTTTAACATCAATGATAAAAGAAATAATGGATTCGATGGTGAGATTCAAGAA
CAACATATTTATATAAATGGAAAAGAATATACATCATTTAATGATATTAAACAAATAATAGACAAGACACTAAAC
ATTAAGATTGTTGTAAAAGATTTTGCAAGAAATACAACCGTAAAAGAATTCATTTTAAATAAAGATACGGGAGAG
GTAAGTGAATTAAAACCTCATAGGGTAACTGTGACCATTCAAAATGGAAAAGAAATGAGTTCAACGATAGTGTCG
GAAGAAGATTTTATTTTACCTGTTTATAAGGGTGAATTAGAAAAAGGATACCAATTTGATGGTTGGGAAATTTCTG
GTTTCGAAGGTAAAAAAGACGCTGGCTATGTTATTAATCTATCAAAAGATACCTTTATAAAACCTGTATTCAAGAA
AATAGAGGAGAAAAAGGAGGAAGAAAATAAACCTACTTTTGATGTATCGAAAAAGAAAGATAACCCACAAGTAA
ACCATAGTCAATTAAATGAAAGTCACAGAAAAGAGGATTTACAAAGAGAAGAGCATTCACAAAAATCTGATTCAA
CTAAGGATGTTACAGCTACAGTTCTTGATAAAAACAATATCAGTAGTAAATCAACTACTAACAATCCTAATAAGTT
GCCAAAAACTGGAACAGCAAGCGGAGCCCAGACACTATTAGCTGCCGGAATAATGTTTATAGTAGGAATTTTTCT
TGGATTGAAGAAAAAAAATCAAGATTAA

YPVVLADTSSSEDALNISDKEKVAENKEKHENIHSAMETSQDFKEKKTAVIKEKEVVSKNPVIDNNTSNEEAKIKEENSN
KSQGDYTDSFVNKNTENPKKEDKVVYIAEFKDKESGEKAIKELSSLKNTKVLYTYDRIFNGSAIETTPDNLDKIKQIEGIS
SVERAQKVQPMMNHARKEIGVEEAIDYLKSINAPFGKNFDGRGMVISNIDTGTDYRHKAMRIDDDAKASMRFKKEDLK
GTDKNYWLSDKIPHAFNYYNGGKITVEKYDDGRDYFDPHGMHIAGILAGNDTEQDIKNFNGIDGIAPNAQIFSYKMYS
DAGSGFAGDETMFHAIEDSIKHNVDVVSVSSGFTGTGLVGEKYWQAIRALRKAGIPMVVATGNYATSASSSSWDLVAN
NHLKMTDTGNVTRTAAHEDAIAVASAKNQTVEFDKVNIGGESFKYRNIGAFFDKSKITTNEDGTKAPSKLKFVYIGKGQ
DQDLIGLDLRGKIAVMDRIYTKDLKNAFKKAMDKGARAIMVVNTVNYYNRDNWTELPAMGYEADEGTKSQVFSISGD
DGVKLWNMINPDKKTEVKRNNKEDFKDKLEQYYPIDMESFNSNKPNVGDEKEIDFKFAPDTDKELYKEDIIVPAGSTS
WGPRIDLLLKPDVSAPGKNIKSTLNVINGKSTYGYMSGTSMATPIVAASTVLIRPKLKEMLERPVLKNLKGDDKIDLTSL
TKIALQNTARPMMDATSWKEKSQYFASPRQQGAGLINVANALRNEVVATFKNTDSKGLVNSYGSISLKEIKGDKKYFTI
KLHNTSNRPLTFKVSASAITTDSLTDRLKLDETYKDEKSPDGKQIVPEIHPEKVKGANITFEHDTFTIGANSSFDLNAVIN
VGEAKNKNKFVESFIHFESVEAMEALNSSGKKINFQPSLSMPLMGFAGNWNHEPILDKWAWEEGSRSKTLGGYDDDG
KPKIPGTLNKGIGGEHGIDKFNPAGVIQNRKDKNTTSLDQNPELFAFNNEGINAPSSSGSKIANIYPLDSNGNPQDAQLER
GLTPSPLVLRSAEEGLISIVNTNKEGENQRDLKVISREHFIRGILNSKSNDAKGIKSSKLKVWGDLKWDGLIYNPRGREEN
APESKDNQDPATKIRGQFEPIAEGQYFYKFKYRLTKDYPWQVSYIPVKIDNTAPKIVSVDFSNPEKIKLITKDTYHKVKD
QYKNETLFARDQKEHPEKFDEIANEVWYAGAALVNEDGEVEKNLEVTYAGEGQGRNRKLDKDGNTIYEIKGAGDLRG
KIIEVIALDGSSNFTKIHRIKFANQADEKGMISYYLVDPDQDSSKYQKLGEIAESKFKNLGNGKEGSLKKDTTGVEHHHQ
ENEESIKEKSSFTIDRNISTIRDFENKDLKKLIKKKFREVDDFTSETGKRMEEYDYKYDDKGNIIAYDDGTDLEYETEKLD

EIKSKIYGVLSPSKDGHFEILGKISNVSKNAKVYYGNNYKSIEIKATKYDFHSKTMTFDLYANINDIVDGLAFAGDMRLF
VKDNDQKKAEIKIRMPEKIKETKSEYPYVSSYGNVIELGEGDLSKNKPDNLTKMESGKIYSDSEKQQYLLKDNIILRKGY
ALKVTTYNPGKTDMLEGNGVYSKEDIAKIQKANPNLRALSETTIYADSRNVEDGRSTQSVLMSALDGFNIIRYQVFTFK
MNDKGEAIDKDGNLVTDSSKLVLFGKDDKEYTGEDKFNVEAIKEDGSMLFIDTKPVNLSMDKNYFNPSKSNKIYVRNP
EFYLRGKISDKGGFNWELRVNESVVDNYLIYGDLHIDNTRDFNIKLNVKDGDIMDWGMKDYKANGFPDKVTDMDGN
VYLQTGYSDLNAKAVGVHYQFLYDNVKPEVNIDPKGNTSIEYADGKSVVFNINDKRNNGFDGEIQEQHIYINGKEYTSF
NDIKQIIDKTLNIKIVVKDFARNTTVKEFILNKDTGEVSELKPHRVTVTIQNGKEMSSTIVSEEDFILPVYKGELEKGYQFD
GWEISGFEGKKDAGYVINLSKDTFIKPVFKKIEEKKEEENKPTFDVSKKKDNPQVNHSQLNESHRKEDLQPEEHSQKSDS
TKDVTATVLDKNNISSKSTTNNPNKLPKTGTASGAQTLLAAGIMFIVGIFLGLKKKNQDZ

ID6 597 bp

CTTGAATTAAATAAAAAACGTCATGCGACTAAGCATTTTACTGATAAGCTTGTTGATCCCAAAGATGTGCGTACGG
CTATCGAAATTGCAACCTTAGCGCCAAGCGCCCACAACAGCCAGCCTTGGAAATTTGTGGTGGTACGTGAGAAAA
ATGCTGAACTGGCAAAGTTAGCTTATGGTTCCAATTTTGAACAGGTATCATCAGCGCCTGTAACCATTGCCTTGTT
TACAGATACGGACTTAGCCAAACGTGCTCGTAAGATTGCCCGTGTTGGTGGTGCTAATAACTTTTCTGAAGAGCAA
CTTCAATATTTTATGAAAAATCTGCCAGCTGAGTTTGCCCGTTACAGTGAGCAACAAGTCAGCGACTACCTAGCTC
TCAATGCAGGTTTGGTTGCCATGAACTTGGTTCTTGCATTGACAGACCAAGGAATTGGTTCTAACATTATTCTTGG
TTTTGACAAATCAAAAGTTAATGAAGTTTTGGAAATCGAAGACCGTTTCCGCCCAGAACTCTTGATCACAGTGGGT
TATACAGACGAAAAATTGGAACCAAGCTACCGCTTGCCAGTAGATGAAATCATCGAGAAAAGATAG

LELNKKRHATKHFTDKLVDPKDVRTAIEIATLAPSAHNSQPWKFVVVREKNAELAKLAYGSNFEQVSSAPVTIALFTDT
DLAKRARKIARVGGANNFSEEQLQYFMKNLPAEFARYSEQQVSDYLALNAGLVAMNLVLALTDQGIGSNIILGFDKSK
VNEVLEIEDRFRPELLITVGYTDEKLEPSYRLPVDEIIEKRZ

ID7 1401 bp

ATGACAGCAATTGATTTTACAGCAGAAGTAGAAAAACGCAAAGAAGACCTCTTGGCTGACTTGTTTAGCCTTTTGG
AAATCAATTCAGAACGTGATGACAGCAAGGCTGATGCCCAGCATCCATTTGGGCCTGGTCCAGTAAAAGCCTTGG
AGAAATTCCTTGAAATCGCAGACCGCGATGGCTACCCAACTAAGAATGTTGATAACTATGCAGGACATTTTGAGTT
TGGTGATGGAGAAGAAGTTCTCGGAATCTTTGCCCATATGGATGTGGTGCCTGCTGGTAGCGGTTGGGACACAGA
CCCTTACACACCAACTATCAAAGATGGTCGCCTTTATGCGCGCGGGGCTTCGGACGATAAGGGTCCTACAACAGC
TTGTTACTATGGTTTGAAAATCATCAAAGAATTGGGTCTTCCAACTTCTAAGAAAGTTCGCTTCATCGTTGGAACA
GACGAAGAATCAGGCTGGGCAGACATGGACTACTACTTTGAGCACGTAGGACTTGCCAAACCAGATTTCGGTTTC
TCACCAGATGCTGAATTTCCAATCATCAATGGTGAAAAAGGAAATATCACGGAATACCTCCACTTTGCAGGAGAA
AATACAGGTGTTGCCCGTCTTCACAGCTTTACAGGTGGTTTACGTGAAAATATGGTACCAGAATCAGCAACAGCA
GTCGTTTCAGGTGACTTGGCTGACTTGCAAGCTAAACTAGATGCCTTTGTTGCAGAACACAAACTTAGAGGAGAAC
TCCAAGAAGAAGCTGGCAAATACAAGGTGACGATCATTGGTAAATCAGCCCACGGTGCTATGCCTGCTTCAGGTG
TCAATGGCGCAACTTACCTTGCCCTCTTCCTCAGCCAGTTTGGCTTTGCTGGTCCAGCCAAAGACTACCTTGACAT
CGCAGGTAAAATTCTCTTGAACGATCATGAGGGTGAAAATCTTAAGATTGCTCATGTGGATGAAAAGATGGGTGC
TCTTTCTATGAATGCCGGCGTCTTCCACTTCGATGAAACAAGTGCTGATAATACCATTGCCCTCAACATCCGCTAT
CCAAAAGGAACAAGTCCAGAACAAATCAAGTCAATCCTTGAAAACTTGCCAGTTGTTTCTGTTAGCCTGTCTGAAC
ACGGTCACACGCCTCACTATGTGCCAATGGAAGATCCACTTGTGCAAACCTTGTTGAATATCTATGAAAAACAAAC
TGGCTTTAAAGGTCATGAACAAGTCATCGGTGGTGGAACCTTTGGTCGCTTGCTAGAACGCGGAGTTGCCTACGGT
GCTATGTTCCCAGACTCGATTGATACCATGCACCAAGCCAATGAATTTATCGCCTTGGATGATCTTTTCCGAGCAG
CAGCAATTTATGCCGAAGCTATTTACGAATTGATCAAATAA

MTAIDFTAEVEKRKEDLLADLFSLLEINSERDDSKADAQHPFGPGPVKALEKFLEIADRDGYPTKNVDNYAGHFEFGDG
EEVLGIFAHMDVVPAGSGWDTDPYTPTIKDGRLYARGASDDKGPTTACYYGLKIIKELGLPTSKKVRFIVGTDEESGWA
DMDYYFEHVGLAKPDFGFSPDAEFPIINGEKGNITEYLHFAGENTGVARLHSFTGGLRENMVPESATAVVSGDLADLQ
AKLDAFVAEHKLRGELQEEAGKYKVTIIGKSAHGAMPASGVNGATYLALFLSQFGFAGPAKDYLDIAGKILLNDHEGE
NLKIAHVDEKMGALSMNAGVFHFDETSADNTIALNIRYPKGTSPEQIKSILENLPVVSVSLSEHGHTPHYVPMEDPLVQT
LLNIYEKQTGFKGHEQVIGGGTFGRLLERGVAYGAMFPDSIDTMHQANEFIALDDLFRAAAIYAEAIYELIKZ

ID8 1617 bp

GTGTATACTATTATAAAATCAAATATAAAAAAATTTAGTTTATTAACGATATTTATTGTTGCTGGTCAATTATTGCT
AATTTATGCAGCAACTATTAATGCTCTGGTGTTGAATGAATTAATTGCGATGAATTTAGAGCGGTTTTTGAAATTG
TCAATCTACCAAATGATTGTCTGGTGTGGGATAATATTCCTTGACTGGGTAGTGAAAAATTATCAGGTTGAAGTGA
TCCAAGAGTTTAATCTAGAGATTCGAAATAGAGTTGCCACAGACATCTCTAACTCTACCTATCAAGAATTTCATAG
TAAATCATCAGGAACATATCTTTCGTGGCTAAATAATGATGTTCAGACTTTAAATGATCAGGCGTTTAAACAACTT
TTTTTAGTAATAAAAGGAATTTCTGGTACTATATTTGCAGTTGTGACTCTTAATCACTATCATTGGTCATTGACTGT
AGCCACCTTGTTTTCATTAATGATTATGCTACTTGTACCAAAAATCTTTGCATCGAAAATGCGAGAAGTTAGTCTA
AATTTAACTAACCAAAATGAAGCTTTTTTAAAATCTAGTGAGACTATATTGAATGGATTTGATGTGTTAGCGTCCT

TGAATCTTTTATATGTATTGCCTAAGAAAATTAAAGAAGCAGGAATTTTATTAAAGATGGTTATACAAAGAAAGAC
AACTGTAGAAACGTTAGCAGGCGCTATTAGCTTCTTTCTCAATATTTTTTTCAGATATCTCTCGTTTTTTAACAG
GCTATCTTGCAATAAAAGGAATAGTGAAAATTGGTACTATTGAAGCAATAGGAGCACTAACAGGTGTTATTTTTAC
AGCGCTAGGTGAATTAGGAGGTCAATTATCCTCTATTATTGGTACGAAGCCTATTTTTTTAAAATTGTATTCAATTA
ATCCAATTGAGTCAAATAAAATGAATGATATCGAACCAAATGAGGTGAATAGAGATTTTCCGTTATATGAAGCAA
AAAATATTTGCTATAAGTATGGAGATAAAGAAATATTAAAAAACTTAAATTTTTGTTTTCAACGTAATGAAAAGTA
TTTAATTTTAGGTGAAAGTGGAAGCGGGAAATCTACATTATTAAAATTATTGAATGGCTTTTTGAGAGATTATAGT
GGAGAATTGCGATTCTGCGGGGATGATATAAAAAAAACCTCCTATTTAAATATGGTTTCGAATGTTCTATATGTAG
ATCAAAAAGCTTATTTGTTTGAAGGTACGATTAGAGATAATATTTTATTGGAAGAAAATTATACTGATGAAGAAAT
ACTACAGTCTTTAGAGCAAGTTGGTTTGAGTGTAAAAGATTTTCCTAATAACATTTTAGATTATTATGTTGGTGATG
ATGGGAGATTACTGTCAGGAGGGCAGAAACAAAAAATTACTTTAGCTAGAGGGCTAATTAGAAATAAGAAAATAG
TATTAATTGACGAGGGAACTTCTGCTATCGATAGGAGAACTTCGTTAGCGATTGAACGTAAGATATTAGATAGAGA
GGATTTGACTGTCATTATTGTTACCCATGCTCCGCATCCGGAACTTAAACAATATTTTACTAAGATATATCAATTTC
CAAAGGATTTTATTTAA

MYTIIKSNIKKFSLLTIFIVAGQLLLIYAATINALVLNELIAMNLERFLKLSIYQMIVWCGIIFLDWVVKNYQVEVIQEFNL
EIRNRVATDISNSTYQEFHSKSSGTYLSWLNNDVQTLNDQAFKQLFLVIKGISGTIFAVVTLNHYHWSLTVATLFSLMIM
LLVPKIFASKMREVSLNLTNQNEAFLKSSETILNGFDVLASLNLLYVLPKKIKEAGILLKMVIQRKTTVETLAGAISFFLNI
FFQISLVFLTGYLAIKGIVKIGTIEAIGALTGVIFTALGELGGQLSSIIGTKPIFLKLYSINPIESNKMNDIEPNEVNRDFPLYE
AKNICYKYGDKEILKNLNFCFQRNEKYLILGESGSGKSTLLKLLNGFLRDYSGELRFCGDDIKKTSYLNMVSNVLYVDQ
KAYLFEGTIRDNILLEENYTDEEILQSLEQVGLSVKDFPNNILDYYVGDDGRLLSGGQKQKITLARGLIRNKKIVLIDEGT
SAIDRRTSLAIERKILDREDLTVIIVTHAPHPELKQYFTKIYQFPKDFIZ

ID9 705 bp

ATAACAGTTAAACAGATTATGGACGAAATAGCCGTTTCAGATATGACTGCAAGGCGCTATTTACAGGAATTAGCT
GATAAAGATTTGCTGATTCGTGTGCATGGTGGAGCTGAAAAACTTCGAACCAACTCCCTTTTGACTAATGAGCGAT
CAAATATTGAAAAACAAGCCCTCCAAACGGCAGAAAAACAAGAAATAGCCCATTTTGCAGGCAGTCTAGTAGAAG
AAAGAGAAACTATTTTCATTGGACCAGGAACAACATTAGAGTTTTTTGCGCGTGAGTTGCCTATTGACAATATCCG
CGTCGTAACCAACAGTCTACCTGTTTTTCTGATTTTAAGCGAACGAAAATTAACAGATTTGATTTTAATAGGTGGA
AATTATCGCGATATTACAGGTGCTTTTGTTGGTACATTGACCCTACAAAATCTCTCTAATCTCCAATTTTCTAAAGC
TTTCGTTAGCTGTAATGGTATTCAAAACGGAGCTCTAGCTACTTTTAGCGAGGAAGAGGGAGAGGCTCAACGCATC
GCTTTAAATAATTCTAATAAAAAATATTTACTCGCAGATCATAGCAAGTTCAATAAGTTTGATTTTTATACTTTTTA
TAATGTATCAAATCTTGATACTATTGTTTCAGATTCTAAACTAAGTGATTCAATCCTTTTTAAGCTATCTAAACACA
TTAAAGTCATCAAGCCTTAA

ITVKQIMDEIAVSDMTARRYLQELADKDLLIRVHGGAEKLRTNSLLTNERSNIEKQALQTAEKQEIAHFAGSLVEERETI
FIGPGTTLEFFARELPIDNIRVVTNSLPVFLILSERKLTDLILIGGNYRDITGAFVGTLTLQNLSNLQFSKAFVSCNGIQNGA
LATFSEEEGEAQRIALNNSNKKYLLADHSKFNKFDFYTFYNVSNLDTIVSDSKLSDSILFKLSKHIKVIKPZ

ID10 483 bp

ATGACTGAGTTTTCGTTAGATCTTCTTCTAGAAGCCATTAAACTAGCTCGTTGGACCTACTACTATCACTTGAAAC
AGCTAGACAAAACAGATAAAGACCAAGAGCTTAAAACTGAAATTCAATCCATCTTTATCGAACACAAGGGAAATT
ATGCTTATCGCCGGGTTCATTTAGAACTAAGAAATCGTGGTTATCTGGTAAATCATAAAAGAGTTCAAGGCTTGaT
GAAAGTACTCAATTTACAAGCTAAAATGCGAAAGAAACGAAAATATTCTTCTCATAAAGGAGACGTTGGTAAGAA
GGCAGAGAATCTCATTCAAGCCCAATTTGAAGGCTCTAAAACAATGGAAAAGTGCTACACAGATGTGACTGAATT
TGCCATTCCAGCAAGTACTCAAAAGCTTTACTTATCACCAGTTTTAGATGGCTTTAACAGCGAAATTATTGCTTTTA
ATCTTTCTTGTTCGCCTAATTTAGAATAA

MTEFSLDLLLEAIKLARWTYYYHLKQLDKTDKDQELKTEIQSIFIEHKGNYAYRRVHLELRNRGYLVNHKRVQGLMKV
LNLQAKMRKKRKYSSHKGDVGKKAENLIQAQFEGSKTMEKCYTDVTEFAIPASTQKLYLSPVLDGFNSEIIAFNLSCSPN
LEZ

ID14 1266 bp

CCAGGATTTGGTACCGTTGCAAGTGGTGTGCCTTTCCTCCTAAAGGAAAATGGAGGAAAAAATCAATCAATCAGCA
CATTCAGATATCAAAGTTGCTAAGGTATTGGTCAAGGATGAAGATGAAAAAAATCGCTTGCTTGCAGCAGGGAAT
GACTTTAACTTTGTAACCAATGTGGATGATATTTTATCAGACCAGGATATTACTATCGTAGTGGAATTGATGGGGC
GTATTGAGCCTGCTAAAACCTTTATCACTCGTGCCTTGGAAGCTGGAAAACACGTTGTTACTGCTAACAAGGACCT
TTTAGCTGTCCATGGCGCAGAATTGCTAGAAATCGCTCAAGCTAACAAGGTAGCACTTTACTACGAAGCAGCAGTT
GCTGGTGGGATTCCAATTCTTCGTACTTTAGCAAATTCCTTGGCTTCTGATAAAATTACGCGCGTGCTTGGAGTAG
TCAACGGAACTTCCAACTTCATGGTGACCAAGATGGTGGAAGAAGGCTGGTCTTACGATGATGCTCTTGCGGAAG
CACAACGTCTAGGATTTGCAGAAAGCGATCCGACGAATGACGTAGATGGGATTGATGCAGCCTACAAGATGGTTA

TTTTGAGCCAATTTGCCTTTGGCATGAAGATTGCCTTTGATGATGTAGCCCACAAGGGAATCCGCAATATCACACC
AGAAGACGTAGCTGTAGCTCAAGAGCTTGGTTACGTAGTGAAATTGGTTGGTTCTATTGAGGAAACTTCTTCAGGT
ATTGCTGCAGAAGTGACTCCAACCTTCCTACCTAAAGCGCACCCACTTGCTAGTGTGAATGGCGTAATGAACGCTG
TCTTTGTAGAATCTATCGGTATTGGTGAGTCTATGTACTACGGACCAGGTGCGGGTCAAAAACCAACTGCAACAAG
TGTTGTAGCTGATATTGTCCGTATCGTTCGTCGTTTGAATGATGGTACTATTGGCAAAGACTTCAACGAATATAGC
CGTGACTTGGTCTTGGCAAATCCTGAAGATGTCAAAGCAAACTACTATTTCTCAATCTTGGCTCTAGACTCAAAAG
GTCAGGTCTTGAAGTTGGCTGAAATCTTCAATGCTCAAGATATTTCCTTTAAGCAAATCCTTCAAGATGGCAAAGA
GGGTGACAAGGCGCGTGTCGTTATCATCACACACAAGATTAATAAAGCCCAGCTTGAAAATGTCTCAGCTGAATT
GAAGAAGGTTTCAGAATTCGACCTCTTGAATACCTTCAAGGTGCTAGGAGAATAA

PGFGTVASGVPFLLKENGGKINQSAHSDIKVAKVLVKDEDEKNRLLAAGNDFNFVTNVDDILSDQDITIVVELMGRIEP
AKTFITRALEAGKHVVTANKDLLAVHGAELLEIAQANKVALYYEAAVAGGIPILRTLANSLASDKITRVLGVVNGTSNF
MVTKMVEEGWSYDDALAEAQRLGFAESDPTNDVDGIDAAYKMVILSQFAFGMKIAFDDVAHKGIRNITPEDVAVAQE
LGYVVKLVGSIEETSSGIAAEVTPTFLPKAHPLASVNGVMNAVFVESIGIGESMYYGPGAGQKPTATSVVADIVRIVRRL
NDGTIGKDFNEYSRDLVLANPEDVKANYYFSILALDSKGQVLKLAEIFNAQDISFKQILQDGKEGDKARVVIITHKINKA
QLENVSAELKKVSEFDLLNTFKVLGEZ

ID16 1725 bp

ATGAAACACCTATTATCTTACTTCAAACCCTACATCAAGGAATCAATTTTAGCCCCCTTGTTCAAGCTGTTAGAAG
CTGTTTTTGAGCTCTTGGTTCCCATGGTGATTGCTGGGATTGTTGACCAATCTTTACCTCAGGGAGATCAAGGTCAT
CTCTGGATGCAGATTGGCCTGCTCCTTATCTTTGCAGTAATTGGCGTTTTAGTGGCCTTGATAGCTCAATTTTACTC
AGCAAAGGCAGCAGTAGGTTCTGCTAAGGAATTGACAAACGATCTTTATCGTCATATTCTTTCCTTGCCCAAGGAC
AGCAGAGACCGTCTGACAACTTCTAGTTTGGTCACTCGCTTGACTTCGGATACCTACCAGATTCAGACTGGTATCA
ATCAATTCCTGCGTCTCTTTTTACGAGCGCCCATTATCGTTTTTGGTGCCATTTTTATGGCTTATCGAATCTCAGCT
GAGTTGACTTTCTGGTTCTTAGTCTTGGTTGCCATTTTGACCATTGTCATTGTAGGGTTATCTCGATTGGTCAATCC
TTTCTACAGTAGTCTCAGAAAGAAAACGGACCAACTGGTTCAGGAAACGCGCCAGCAATTGCAAGGGATGCGGGT
TATTCGTGCTTTTGGTCAAGAAAAACGAGAGTTACAGATTTTTCAAACCCTTAACCAAGTTTATGCTAGATTACAA
GAAAAGACAGGTTTCTGGTCTAGTTTATTAACACCTCTGACCTATCTGATTGTCAATGGAACTCTTCTCGTTATTAT
CTGGCAAGGCTATATTTCAATTCAAGGAGGAGTGCTCAGTCAAGGTGCTCTCATTGCTCTTATCAATTACCTCTTA
CAGATTTTGGTGGAATTGGTCAAGCTAGCCATGTTGATCAATTCCCTCAACCAGTCCTATATCTCAGTCAAGCGAA
TCGAGGAAGTCTTTGTTGAGGCTCCAGAGGATATCCATTCAGAGTTAGAACAAAAGCAAGCTACCAGAGATAAGG
TTTTACAAGTCCAAGAATTGACCTTTACCTATCCTGATGCGGCCCAGCCTTCTCTGAGATACATTTCCTTTGATATG
ACTCAAGGACAAATTCTAGGTATCATCGGGGGAACTGGTTCTGGTAAATCAAGCTTGGTGCAACTCTTACTTGGAC
TTTATCCAGTAGACAAGGGGAACATTGACCTTTATCAAAATGGACGTAGTCCTCTTAATTTGGAGCAGTGGCGGTC
TTGGATTGCCTATGTACCTCAAAAGGTCGAACTCTTTAAAGGAACCATTCGTTCCAACTTGACTCTAGGTTTCAAT
CAAGAAGTATCTGACCAGGAACTCTGGCAGGCCTTGGAGATTGCGCAAGCTAAGGATTTTGTCAGTGAAAAGGAA
GGACTCTTGGATGCTCTAGTTGAGGCAGGGGGGCGAAATTTCTCAGGTGGACAAAAACAAAGATTGTCTATCGCC
CGAGCAGTCTTGCGCCAGGCTCCGTTTCTCATCCTAGATGATGCAACCTCGGCACTGGATACCATTACAGAGTCCA
AGCTCTTGAAAGCTATTAGAGAAAATTTTCCAAACACGAGCTTAATTTTGATCTCTCAACGAACCTCAACTTTACA
GATGGCGGACCAGATTCTCCTCTTGGAAAAAGGTGAGTTGCTAGCTGTTGGCAAGCACGATGACTTGATGAAATC
CAGCCAAGTCTATTGTGAAATCAATGCATCCCAACATGGAAAGGAGGACTAG

MKHLLSYFKPYIKESILAPLFKLLEAVFELLVPMVIAGIVDQSLPQGDQGHLWMQIGLLLIFAVIGVLVALIAQFYSAKAA
VGSAKELTNDLYRHILSLPKDSRDRLTTSSLVTRLTSDTYQIQTGINQFLRLFLRAPIIVFGAIFMAYRISAELTFWFLVLV
AILTIVIVGLSRLVNPFYSSLRKKTDQLVQETRQQLQGMRVIRAFGQEKRELQIFQTLNQVYARLQEKTGFWSSLLTPLT
YLIVNGTLLVIIWQGYISIQGGVLSQGALIALINYLLQILVELVKLAMLINSLNQSYISVKRIEEVFVEAPEDIHSELEQKQA
TRDKVLQVQELTFTYPDAAQPSLRYISFDMTQGQILGIIGGTGSGKSSLVQLLLGLYPVDKGNIDLYQNGRSPLNLEQWR
SWIAYVPQKVELFKGTIRSNLTLGFNQEVSDQELWQALEIAQAKDFVSEKEGLLDALVEAGGRNFSGGQKQRLSIARAV
LRQAPFLILDDATSALDTITESKLLKAIRENFPNTSLILISQRTSTLQMADQILLLEKGELLAVGKHDDLMKSSQVYCEINA
SQHGKEDZ

ID18 1224 bp

ATGAAACGTTCTCTCGACTCAAGAGTCGATTACAGTTTGCTCTTGCCAGTATTTTTTCTACTGGTCATCGGTGTGGT
GGCTATCTATATAGCCGTTAGTCATGATTATCCCAATAATATTCTGCCCATTTTAGGGCAGCAGGTCGCCTGGATT
GCCTTGGGGCTTGTGATTGGTTTTGTGGTCATGCTCTTTAATACAGAATTTCTTTGGAAGGTGACCCCCTTTCTATA
TATTTTAGGCTTGGGACTTATGATCTTGCCGATTGTATTTTATAATCCAAGCTTAGTTGCATCAACGGGTGCCAAA
AACTGGGTATCAATAAATGGAATTACCCTATTCCAACCGTCAGAATTTATGAAGATATCCTATATCCTCATGTTGG
CTCGTGTCATTGTCCAATTTACAAAGAAACATAAGGAATGGAGACGCACGGTTCCGCTGGACTTTTTGTTAATTTT
CTGGATGATTCTCTTTACCATTCCAGTCCTAGTTCTTTTAGCACTTCAAAGTGACTTGGGGACGGCTTTGGTTTTTG
TAGCCATTTTTCTCAGGAATCGTTTTATTATCAGGGGTTTCTTGGAAAATTATTATCCCAGTATTTGTGACTGCTGTA
ACAGGAGTTGCTGGTTTCTTAGCTATCTTTATTAGCAAGGACGGACGAGCTTTTCTTCACCAGATTGGAATGCCGA
CCTACCAAATTAATCGGATTTTGGCTTGGCTCAATCCCTTTGAGTTTGCCCAAACAACGACTTACCAGCAGGCTCA

AGGGCAGATTGCCATTGGGAGTGGTGGCTTATTTGGTCAGGGATTTAATGCTTCGAATCTGCTTATCCCAGTTCGA
GAGTCAGATATGATTTTTACGGTTATTGCAGAAGATTTTGGCTTTATTGGCTCTGTCCTGGTTATTGCCCTCTATCT
CATGTTGATTTACCGTATGTTGAAGATTACTCTTAAATCAAATAACCAGTTCTACACTTATATTTCCACAGGTTTGA
TTATGATGTTGCTCTTCCACATCTTTGAGAATATCGGTGCTGTGACTGGACTACTTCCTTTGACGGGGATTCCCTTG
CCTTTCATTTCGCAAGGGGGATCAGCTATTATCAGTAATCTGATTGGTGTTGGTTTGCTTTTATCGATGAGTTACCA
GACTAATCTAGCTGAAGAAAAGAGCGGAAAAGTCCCATTCAAACGGAAAAAGGTTGTATTAAAACAAATTAAATA
A

MKRSLDSRVDYSLLLPVFFLLVIGVVAIYIAVSHDYPNNILPILGQQVAWIALGLVIGFVVMLFNTEFLWKVTPFLYILGL
GLMILPIVFYNPSLVASTGAKNWVSINGITLFQPSEFMKISYILMLARVIVQFTKKHKEWRRTVPLDFLLIFWMILFTIPVL
VLLALQSDLGTALVFVAIFSGIVLLSGVSWKIIIPVFVTAVTGVAGFLAIFISKDGRAFLHQIGMPTYQINRILAWLNPFEF
AQTTTYQQAQGQIAIGSGGLFGQGFNASNLLIPVRESDMIFTVIAEDFGFIGSVLVIALYLMLIYRMLKITLKSNNQFYTYI
STGLIMMLLFHIFENIGAVTGLLPLTGIPLPFISQGGSAIISNLIGVGLLLSMSYQTNLAEEKSGKVPFKRKKVVLKQIKZ

ID22 987 bp

ATGGTGGCTAAGAAAAAAATCTTATTTTTTATGTGGTCTTTTTCTCTTGGAGGTGGTGCAGAGAAGATTCTATCAA
CCATTGTTTCAAATCTGGATCCAGAAAAGTATGATATTGATATTCTTGAAATGGAGCACTTTGACAAGGGATATGA
ATCTGTTCCAAAGCATGTACGCATTTTAAAATCCCTTCAAGATTATCGCCAAACCAGATGGTTACGAGCTTTTTTG
TGGAGAATGAGAATTTATTTTCCAAGACTGACTCGTCGTTTGCTTGTAAAAGATGATTATGATGTTGAAGTTTCTTT
TACCATTATGAATCCACCACTGTTGTTCTCTAAAAGAAGAGAAGTCAAGAAGATATCTTGGATTCATGGAAGTATT
GAAGAACTTCTTAAGGATAGCTCTAAAAGAGAATCACATAGAAGCCAGTTGGATGCTGCGAATACAATTGTAGGG
ATTTCAAAAAAGACCAGCAATTCTATCAAGGAAGTTTATCCAGATTATACTTCTAAATTACAGACAATCTACAATG
GATATGATTTTCAGACTATTCTAGAAAAATCTCAAGAGAAGATCGATATCGAGATTGCTCCTCAAAGTATCTGTAC
TATCGGACGGATTGAGGAAAATAAGGGTTCTGACCGTGTAGTGGAAGTGATACGATTATTACACCAAGAGGGAAA
AAACTATCATCTCTATTTTATCGGGGCTGGTGATATGGAAGAGGAACTGAAAAAACGAGTCAAAGAGTATGGGAT
TGAGGACTATGTACATTTCCTTGGTTATCAAAAAAATCCTTATCAGTATCTATCTCAGACGAAAGTTCTTTTGTCTA
TGTCTAAACAAGAAGGTTTTCCTGGAGTGTATGTGGAGGCCTTGAGTCTGGGACTCCCTTTTATCTCTACGGACGT
TGGAGGGGCTGAGGAATTATCCCAAGAAGGACGATTTGGACAAATCATTGAGAGCAATCAAGAGGCAGCTCAGGC
GATTACTAATTACATGACTTCTGCCTCAAACTTTGATGTCGATGAGGCTAGCCAATTCATTCAACAATTTACAATT
ACAAAACAAATCGAACAAGTAGAAAAACTATTAGAGGAGTAG

MVAKKKILFFMWSFSLGGGAEKILSTIVSNLDPEKYDIDILEMEHFDKGYESVPKHVRILKSLQDYRQTRWLRAFLWRM
RIYFPRLTRRLLVKDDYDVEVSFTIMNPPLLFSKRREVKKISWIHGSIEELLKDSSKRESHRSQLDAANTIVGISKKTSNSIK
EVYPDYTSKLQTIYNGYDFQTILEKSQEKIDIEIAPQSICTIGRIEENKGSDRVVEVIRLLHQEGKNYHLYFIGAGDMEEEL
KKRVKEYGIEDYVHFLGYQKNPYQYLSQTKVLLSMSKQEGFPGVYVEALSLGLPFISTDVGGAEELSQEGRFGQIIESNQ
EAAQAITNYMTSASNFDVDEASQFIQQFTITKQIEQVEKLLEEZ

ID23 1434 bp

ATGGAAACTGCATTAATTAGTGTGATTGTGCCAGTCTATAATGTGGCGCAGTACCTAGAAAAATCGATAGCTTCCA
TTCAGAAGCAGACCTATCAAAATCTGGAAATTATTCTTGTTGATGATGGTGCAACAGATGAAAGTGGTCGCTTGTG
TGATTCAATCGCTGAACAAGATGACAGGGTGTCAGTGCTTCATAAAAAGAACGAAGGATTGTCGCAAGCACGAAA
TGATGGGATGAAGCAGGCTCACGGGGATTATCTGATTTTTATTGACTCAGATGATTATATCCATCCAGAAATGATT
CAGAGCTTATATGAGCAATTAGTTCAAGAAGATGCGGATGTTTCGAGCTGTGGTGTCATGAATGTCTATGCTAATG
ATGAAAGCCCACAGTCAGCCAATCAGGATGACTATTTTGTCTGTGATTCTCAAACATTTCTAAAGGAATACCTCAT
AGGTGAAAAAATACCTGGGACGATTTGCAATAAGCTAATCAAGAGACAGATTGCAACTGCCCTATCCTTTCCTAA
GGGGTTGATTTACGAAGATGCCTATTACCATTTTGATTTAATCAAGTTGGCCAAGAAGTATGTGGTTAATACTAAA
CCCTATTATTACTATTTCCATAGAGGGGATAGTATTACGACCAAACCCTATGCAGAGAAGGATTTAGCCTATATTG
ATATCTACCAAAAGTTTTATAATGAAGTTGTGAAAAACTATCCTGACTTGAAAGAGGTCGCTTTTTTCAGATTGGC
CTATGCCCACTTCTTTATTCTGGATAAGATGTTGCTAGATGATCAGTATAAACAGTTTGAAGCCTATTCTCAGATTC
ATCGTTTTTTAAAAGGCCATGCCTTTGCTATTTCTAGGAATCCAATTTTCCGTAAGGGGAGAAGAATTAGTGCTTT
GGCCCTATTCATAAATATTTCCTTATATCGATTCTTATTACTGAAAAATATTGAAAAATCTAAAAAATTACATTAG

METALISVIVPVYNVAQYLEKSIASIQKQTYQNLEIILVDDGATDESGRLCDSIAEQDDRVSVLHKKNEGLSQARNDGMK
QAHGDYLIFIDSDDYIHPEMIQSLYEQLVQEDADVSSCGVMNVYANDESPQSANQDDYFVCDSQTFLKEYLIGEKIPGTI
CNKLIKRQIATALSFPKGLIYEDAYYHFDLIKLAKKYVVNTKPYYYYFHRGDSITTKPYAEKDLAYIDIYQKFYNEVVKN
YPDLKEVAFFRLAYAHFFILDKMLLDDQYKQFEAYSQIHRFLKGHAFAISRNPIFRKGRRISALALFINISLYRFLLLKNIE
KSKKLHZ

ID24 735bp

ATGAGAATCAAAGAGAAAACCAATAATATTAATGGAGGAATAAAAAATGTAAGTAAGCATTATGGTCATTCAATC
ATTCTCAAAGATATAAATTTTGCACTTAACAAGGGTGAAATTGTTGGTCTAGCAGGGAGAAATGGAGTTGGTAAG
AGTACGTTGATGAAAATTCTTGTTCAGAATAATCAACCGACTTCAGGTAATATTATAAGCAGTGATAATGTTGGGT
ATTTAATCGAAGAACCAAAATTATTTTTATCTAAAACAGGTTTAGAGAATTTAAAATATTTGTCAAATTTATATGG
TGTTGACTACAATCAAGAAAGATTTAGATGTTTGATCCAAGAGTTAGATTTGACTCAGTCTATTAATAAAAAAGTA
AAGACCTATTCTTTGGGTACAAAACAAAAATTAGCTTTGCTTCTAACTCTCGTTACGGAACCTGATATATTGATTTT
AGATGAACCGACTAATGGTTTAGATATTGAATCATCACAAATAGTTTTAGCGGTTCTAAAAAAATTAGCTTTACAT
GAAAATGTGGGAATTTTAATATCGAGTCATAAATTAGAAGACATTGAAGAAATTTGTGAGAGAGTTCTTTTCTTGG
AGAACGGGCTTTTGACATTTCAAAAAGTAGGAAAAGATAGTCATAATTTCTTGTTTGAGATAGCTTTTTCATCAGC
TACAGATAGAGACATTTTCATTACCAAACAAGAATTTTGGGATATTGTTTAG

MRIKEKTNNINGGIKNVSKHYGHSIILKDINFALNKGEIVGLAGRNGVGKSTLMKILVQNNQPTSGNIISSDNVGYLIEEP
KLFLSKTGLENLKYLSNLYGVDYNQERFRCLIQELDLTQSINKKVKTYSLGTKQKLALLLTLVTEPDILILDEPTNGLDIE
SSQIVLAVLKKLALHENVGILISSHKLEDIEEICERVLFLENGLLTFQKVGKDSHNFLFEIAFSSATDRDIFITKQEFWDIVZ

ID25 1704bp

ATGACTGAATTAGATAAACGTCACCGCAGTAGCATTTATGACAGCATGGTTAAATCACCTAACCGTGCTATGCTTC
GTGCGACTGGTATGACAGATAAGGACTTTGAAACATCGATTGTGGGAGTGATTTCGACTTGGGCGGAAAATACAC
CATGTAACATTCACTTGCATGATTTCGGGAAACTGGCTAAAGAAGGTGTCAAATCTGCAGGCGCTTGGCCTGTACA
GTTTGGAACCATTACCGTAGCGGACGGGATCGCTATGGGAACGCCTGGTATGCGTTTCTCTCTAACATCTCGTGAC
ATCATCGCGGACTCCATCGAGGCGGCTATGAGTGGTCACAACGTGGATGCCTTCGTCGCTATCGGTGGCTGTGACA
AGAACATGCCTGGATCTATGATTGCTATTGCTAATATGGATATCCCAGCTATTTTCGCCTATGGTGGAACTATTGC
ACCGGGAAATCTTGATGGTAAAGATATCGACTTGGTTTCTGTCTTTGAAGGTATCGGAAAATGGAACCACGGTGAC
ATGACAGCTGAGGACGTGAAACGTCTTGAATGTAATGCCTGCCCTGGCCCTGGTGGTTGTGGTGGTATGTATACTG
CTAATACCATGGCAACTGCTATCGAAGTTCTAGGGATGAGTTTGCCAGGGTCATCCTCTCACCCAGCTGAATCAGC
TGATAAGAAAGAAGATATCGAAGCAGCAGGACGTGCTGTTGTTAAGATGTTGGAACTTGGTCTCAAACCATCAGA
TATCTTGACTCGTGAAGCCTTTGAAGATGCTATCACTGTAACGATGGCTCTCGGTGGTTCTACAAACGCCACTCTT
CACTTGCTCGCCATTGCCCATGCCGCAAATGTTGACTTGTCACTTGAGGACTTCAATACGATTCAAGAACGTGTGC
CTCACTTGGCCGACTTGAAACCATCTGGTCAGTATGTCTTCCAAGACCTCTACGAAGTCGGTGGTGTCCCTGCGGT
TATGAAGTATTTGTTGGCAAATGGTTTCCTTCACGGAGATCGCATCACATGTACTGGTAAGACTGTAGCTGAAAAC
TTGGCTGACTTTGCAGACTTGACTCCAGGCCAAAAAGTTATCATGCCACTTGAAAATCCAAAACGTGCGGATGGTC
CGCTTATCATCTTGAACGGGAACCTTGCTCCTGACGGTGCAGTTGCCAAGGTATCAGGTGTTAAAGTGCGTCGTCA
CGTTGGGCCAGCTAAGGTCTTTGACTCAGAAGAAGATGCGATTCAGGCCGTTCTGACAGATGAAATCGTTGATGG
CGATGTAGTCGTTGTTCGTTTTGTTGGACCTAAAGGTGGTCCTGGTATGCCTGAGATGCTATCACTTTCTTCAATGA
TTGTTGGTAAAGGTCAGGGAGATAAGGTGGCCCTCTTGACGGACGGACGTTTCTCTGGTGGTACTTATGGTCTGGT
TGTTGGACATATCGCTCCTGAAGCTCAGGATGGTGGACCAATTGCCTATCTCCGTACCGGCGATATCGTTACGGTT
GACCAAGATACCAAAGAAATTTCTATGGCCGTATCCGAAGAAGAACTTGAAAAACGCAAGGCAGAAACAACCTTG
CCACCACTTTACAGCCGTGGTGTCCTCGGTAAATATGCCCACATCGTATCATCTGCTTCACGCGGAGCCGTGACAG
ACTTCTGGAATATGGACAAGTCAGGTAAAAAATAA

MTELDKRHRSSIYDSMVKSPNRAMLRATGMTDKDFETSIVGVISTWAENTPCNIHLHDFGKLAKEGVKSAGAWPVQFG
TITVADGIAMGTPGMRFSLTSRDIIADSIEAAMSGHNVDAFVAIGGCDKNMPGSMIAIANMDIPAIFAYGGTIAPGNLDG
KDIDLVSVFEGIGKWNHGDMTAEDVKRLECNACPGPGGCGGMYTANTMATAIEVLGMSLPGSSSHPAESADKKEDIEA
AGRAVVKMLELGLKPSDILTREAFEDAITVTMALGGSTNATLHLLAIAHAANVDLSLEDFNTIQERVPHLADLKPSGQY
VFQDLYEVGGVPAVMKYLLANGFLHGDRITCTGKTVAENLADFADLTPGQKVIMPLENPKRADGPLIILNGNLAPDGA
VAKVSGVKVRRHVGPAKVFDSEEDAIQAVLTDEIVDGDVVVVRFVGPKGGPGMPEMLSLSSMIVGKGQGDKVALLTD
GRFSGGTYGLVVGHIAPEAQDGGPIAYLRTGDIVTVDQDTKEISMAVSEEELEKRKAETTLPPLYSRGVLGKYAHIVSSA
SRGAVTDFWNMDKSGKKZ

ID26 274bp

ATGTTATAATAAAAATAAAGAATTTAAGGAGAAATACAATATGTCAATTTTTATTGGAGGAGCATGGCCATATGC
AAACGGTTCGTTACATATTGGTCACGCGGCAGCGCTTTTACCGGGGGATATTCTTGCAAGATACTATCGTCAGAAG
GGAGAGGAAGTTTTATATGTTTCTGGAAGTGATTGTAATGGAACCCCTATTTCTATCAGAGCTAAAAAAGAAAAT
AAGTCTGTGAAAGAAATTGCTGATTTTTATCATAAGGAATTTAATCCA

CYNKNKEFKEKYNMSIFIGGAWPYANGSLHIGHAAALLPGDILARYYRQKGEEVLYVSGSDCNGTPISIRAKKENKSVK
EIADFYHKEFNP

ID28 1065bp

ATGACAACATTATTTTCAAAAATTAAAGAAGTAACAGAACTTGCTGCAGTCTCAGGTCATGAAGCGCCTGTCCGTG
CTTATCTTCGTGAAAAGTTGACACCGCATGTGGATGAAGTGGTGACAGATGGCTTGGGTGGTATTTTTGGTATCAA

ACATTCAGAAGCTGTGGATGCACCGCGCGTCTTGGTCGCTTCTCATATGGACGAAGTTGGTTTTATGGTCAGCGAA
ATCAAGCCAGATGGTACCTTCCGTGTCGTAGAAATCGGTGGCTGGAACCCCATGGTGGTTAGCAGCCAACGTTTCA
AACTCTTGACTCGTGATGGTCATGAAATTCCTGTGATTTCAGGTTCTGTTCCTCCGCATTTGACTCGTGGAAAGGG
GGGACCAACCATGCCAGCCATTGCCGATATCGTTTTTGATGGTGGTTTTGCGGACAAGGCTGAGGCAGAAAGTTTT
GGCATCCGTCCTGGTGATACCATTGTACCAGATAGTTCTGCAATTTTGACAGCCAATGAAAAAAATATCATCTCAA
AAGCTTGGGATAACCGCTACGGTGTCCTCATGGTAAGCGAGCTAGCTGAAGCTTTATCGGGTCAAAAACTCGGCA
ATGAACTCTATCTGGGTTCTAACGTCCAAGAAGAAGTTGGTCTGCGTGGCGCTCATACCTCTACAACCAAGTTTGA
CCCAGAAGTCTTCCTCGCAGTTGATTGCTCACCAGCAGGTGATGTCTACGGTGGTCAAGGCAAGATTGGAGATGG
AACCTTGATTCGTTTCTATGATCCAGGTCACTTGCTTCTCCCAGGGATGAAGGATTTCCTTTTGACAACGGCTGAA
GAAGCTGGTATCAAGTACCAATACTACTGTGGTAAAGGCGGAACAGATGCAGGTGCAGCTCATCTGAAAAATGGT
GGTGTCCCATCAACAACTATCGGTGTCTGCGCTCGTTATATCCATTCTCACCAAACCCTCTATGCAATGGATGACT
TCCTAGAAGCGCAAGCTTTCTTACAAGCCTTGGTGAAGAAATTGGATCGTTCAACGGTTGATTTGATTAAACATTA
TTAA

MTTLFSKIKEVTELAAVSGHEAPVRAYLREKLTPHVDEVVTDGLGGIFGIKHSEAVDAPRVLVASHMDEVGFMVSEIKP
DGTFRVVEIGGWNPMVVSSQRFKLLTRDGHEIPVISGSVPPHLTRGKGGPTMPAIADIVFDGGFADKAEAESFGIRPGDTI
VPDSSAILTANEKNIISKAWDNRYGVLMVSELAEALSGQKLGNELYLGSNVQEEVGLRGAHTSTTKFDPEVFLAVDCSP
AGDVYGGQGKIGDGTLIRFYDPGHLLLPGMKDFLLTTAEEAGIKYQYYCGKGGTDAGAAHLKNGGVPSTTIGVCARYI
HSHQTLYAMDDFLEAQAFLQALVKKLDRSTVDLIKHYZ

ID31 1182bp

ATGGAATTTTCTATGAAATCAGTCAAAGGACTACTCTTTATCATAGCTAGTTTTATCTTGACTCTTTTGACTTGGAT
GAACACTTCTCCCCAATTCATGATTCCAGGACTAGCTTTAACAAGCCTATCTCTGACTTTTATCCTAGCCACTCGTC
TCCCACTACTAGAAAGCTGGTTTCACAGTTTGGAGAAGGTCTACACCGTCCACAAATTCACAGCCTTTCTCTCAAT
CATCCTACTAATCTTTCATAACTTTAGTATGGGCGGTTTGTGGGGCTCTCGCTTAGCTGCTCAGTTTGGCAATCTTG
CCATCTATATCTTTGCCAGCATCATCCTTGTCGCCTATTTAGGCAAATACATCCAATACGAAGCTTGGCGATGGAT
TCACCGCCTGGTTTACCTAGCCTATATTTTAGGACTCTTTCACATCTACATGATAATGGGCAATCGTCTCCTTACAT
TTAATCTTCTAAGTTTTCTTGTTGGTAGCTATGCCCTTTTAGGCTTACTAGCTGGTTTTTATATCATTTTTCTATATC
AAAAGATTTCCTTCCCCTATCTAGGGAAAATTACCCATCTCAAACGCTTAAATCACGATACTAGAGAAATTCAAAT
CCATCTTAGCAGACCTTTCAACTATCAATCAGGACAATTTGCCTTTCTAAAGATTTTCCAAGAAGGCTTTGAAAGT
GCTCCGCATCCCTTTTCTATCTCAGGAGGTCATGGTCAAACTCTTTACTTTACTGTTAAAACTTCAGGCGACCATAC
CAAGAATATCTATGATAATCTTCAAGCCGGCAGCAAAGTAACCCTAGACAGAGCTTACGGACACATGATCATAGA
AGAAGGACGAGAAAATCAGGTTTGGATTGCTGGAGGTATTGGGATCACCCCCTTCATCTCTTACATCCGTGAACAT
CCTATTTTAGATAAACAGGTTCACTTCTACTATAGCTTCCGTGGAGATGAAAATGCAGTCTACCTAGATTTACTCC
GTAACTATGCTCAGAAAAATCCTAATTTTGAACTCCATCTAATCGACAGTACGAAAGACGGCTATCTTAATTTTGA
ACAAAAAGAAGTGCCCGAACATGCAACCGTCTATATGTGTGGTCCTATTTCTATGATGAAGGCACTTGCCAAACA
GATTAAGAAACAAAATCCAAAAACAGAGCATATTTAC

MEFSMKSVKGLLFIIASFILTLLTWMNTSPQFMIPGLALTSLSLTFILATRLPLLESWFHSLEKVYTVHKFTAFLSIILLIFH
NFSMGGLWGSRLAAQFGNLAIYIFASIILVAYLGKYIQYEAWRWIHRLVYLAYILGLFHIYMIMGNRLLTFNLLSFLVGS
YALLGLLAGFYIIFLYQKISFPYLGKITHLKRLNHDTREIQIHLSRPFNYQSGQFAFLKIFQEGFESAPHPFSISGGHGQTLY
FTVKTSGDHTKNIYDNLQAGSKVTLDRAYGHMIIEEGRENQVWIAGGIGITPFISYIREHPILDKQVHFYYSFRGDENAV
YLDLLRNYAQKNPNFELHLIDSTKDGYLNFEQKEVPEHATVYMCGPISMMKALAKQIKKQNPKTEHIY

ID32 900bp

ATGACTTTTAAATCAGGCTTTGTAGCCATTTTAGGACGTCCCAATGTTGGGAAGTCAACCTTTTTAAATCACGTTAT
GGGGCAAAAGATTGCCATCATGAGTGACAAGGCGCAGACAACGCGCAATAAAATCATGGGAATTTACACGACTGA
TAAGGAGCAAATTGTCTTTATCGACACACCAGGGATTCACAAGCCTAAAACAGCTCTCGGAGATTTCATGGTTGA
GTCTGCCTACAGTACCCTTCGCGAAGTGGACACTGTTCTTTTCATGGTGCCTGCTGATGAAGCGCGTGGTAAGGGG
GACGATATGATTATCGAGCGTCTCAAGGCTGCCAAGGTTCCTGTGATTTTGGTGGTGAATAAAATCGATAAGGTCC
ATCCAGACCAGCTCTTGTCTCAGATTGATGACTTCCGTAATCAAATGGACTTTAAGGAAATTGTTCCAATCTCAGC
CCTTCAGGGAAATAACGTGTCTCGTCTAGTGGATATTTTGAGTGAAAATCTGGATGAAGGTTTCCAATATTTCCCG
TCTGATCAAATCACAGACCATCCAGAACGTTTCTTGGTTTCAGAAATGGTTCGCGAGAAAGTCTTGCACCTAACTC
GTGAAGAGATTCCGCATTCTGTAGCAGTAGTTGTTGACTCTATGAAACGAGACGAAGAGACAGACAAGGTTCACA
TCCGTGCAACCATCATGGTCGAGCGCGATAGCCAAAAAGGGATTATCATCGGTAAAGGTGGCGCTATGCTTAAGA
AAATCGGTAGCATGGCCCGTCGTGATATCGAACTCATGCTAGGAGACAAGGTCTTCCTAGAAACCTGGGTCAAGG
TCAAGAAAAACTGGCGCGATAAAAAGCTAGATTTGGCTGACTTTGGCTATAATGAAAGAGAATACTAA

MTFKSGFVAILGRPNVGKSTFLNHVMGQKIAIMSDKAQTTRNKIMGIYTTDKEQIVFIDTPGIHKPKTALGDFMVESAYS
TLREVDTVLFMVPADEARGKGDDMIIERLKAAKVPVILVVNKIDKVHPDQLLSQIDDFRNQMDFKEIVPISALQGNNVS
RLVDILSENLDEGFQYFPSDQITDHPERFLVSEMVREKVLHLTREEIPHSVAVVVDSMKRDEETDKVHIRATIMVERDSQ
KGIIIGKGGAMLKKIGSMARRDIELMLGDKVFLETWVKVKKNWRDKKLDLADFGYNEREYZ

ID33 855bp

CTGCTTCTTGTTTTTACAGAAGGAGGACTTATGCCTGAATTACCTGAGGTTGAAACCGTTTGTCGTGGCTTAGAAA
AATTGATTATAGGAAAGAAGATTTCGAGTATAGAAATTCGCTACCCCAAGATGATTAAGACGGATTTGGAAGAGT
TTCAAAGGGAATTGCCTAGTCAGATTATCGAGTCAATGGGACGTCGTGGAAAATATTTGCTTTTTTATCTGACAGA
CAAGGTCTTGATTTCCCATTTGCGGATGGAGGGCAAGTATTTTTACTATCCAGACCAAGGACCTGAACGCAAGCAT
GCCCATGTTTTCTTTCATTTTGAAGATGGTGGCACGCTTGTTTATGAGGATGTTCGCAAGTTTGGAACCATGGAAC
TCTTGGTGCCTGACCTTTTAGACGTCTACTTTATTTCTAAAAAATTAGGTCCTGAACCAAGCGAACAAGACTTTGA
TTTACAGGTCTTTCAATCTGCCCTTGCCAAGTCCAAAAAGCCTATCAAATCCCATCTCCTAGACCAGACCTTGGTA
GCTGGACTTGGCAATATCTATGTGGATGAGGTTCTCTGGCGAGCTCAGGTTCATCCAGCTAGACCTTCCCAGACTT
TGACAGCAGAAGAAGCGACTGCCATTCATGACCAGACCATTGCTGTTTTGGGCCAGGCTGTTGAAAAAGGTGGCT
CCACCATTCGGACTTATACCAATGCCTTTGGGGAAGATGGAAGCATGCAGGACTTTCATCAGGTCTATGATAAGAC
TGGTCAAGAATGTGTACGCTGTGGTACCATCATTGAGAAAATTCAACTAGGCGGACGTGGAACCCACTTTTGTCCA
AACTGTCAAAGGAGGGACTGA

MLLVFTEGGLMPELPEVETVCRGLEKLIIGKKISSIEIRYPKMIKTDLEEFQRELPSQIIESMGRRGKYLLFYLTDKVLISHL
RMEGKYFYYPDQGPERKHAHVFFHFEDGGTLVYEDVRKFGTMELLVPDLLDVYFISKKLGPEPSEQDFDLQVFQSALA
KSKKPIKSHLLDQTLVAGLGNIYVDEVLWRAQVHPARPSQTLTAEEATAIHDQTIAVLGQAVEKGGSTIRTYTNAFGED
GSMQDFHQVYDKTGQECVRCGTIIEKIQLGGRGTHFCPNCQRRDZ

ID34 633bp

TTGTCCAAACTGTCAAAGGAGGGACTGATGGGAAAAATCATCGGAATCACTGGGGGAATTGCCTCTGGTAAGTCA
ACTGTGACAAATTTTCTAAGACAGCAAGGCTTTCAAGTAGTGGATGCCGACGCAGTCGTCCACCAACTACAGAAA
CCTGGTGGTCGTCTGTTTGAGGCTCTAGTACAGCACTTTGGGCAAGAAATCATTCTTGAAAACGGAGAACTCAATC
GCCCTCTCCTAGCTAGTCTCATCTTTTCAAATCCTGATGAACGAGAATGGTCTAAGCAAATTCAAGGGGAGATTAT
CCGTGAGGAACTGGCTACTTTGAGAGAACAGTTGGCTCAGACAGAAGAGATTTTCTTCATGGATATTCCCCTACTT
TTTGAGCAGGACTACAGCGATTGGTTTGCTGAGACTTGGTTGGTCTATGTGGACCGAGATGCCCAAGTGGAACGCT
TAATGAAAAGGGACCAGTTGTCCAAAGATGAAGCTGAGTCTCGTCTGGCAGCCCAGTGGCCTTTAGAAAAAAAGA
AAGATTTGGCCAGCCAGGTTCTTGATAATAATGGCAATCAGAACCAGCTTCTTAATCAAGTGCATATCCTTCTTGA
GGGAGGTAGGCAAGATGACAGAGATTAA

MSKLSKEGLMGKIIGITGGIASGKSTVTNFLRQQGFQVVDADAVVHQLQKPGGRLFEALVQHFGQEIILENGELNRPLLA
SLIFSNPDEREWSKQIQGEIIREELATLREQLAQTEEIFFMDIPLLFEQDYSDWFAETWLVYVDRDAQVERLMKRDQLSK
DEAESRLAAQWPLEKKKDLASQVLDNNGNQNQLLNQVHILLEGGRQDDRDZ

ID35 1269bp

TTGATAATAATGGCAATCAGAACCAGCTTCTTAATCAAGTGCATATCCTTCTTGAGGGAGGTAGGCAAGATGACA
GAGATTAACTGGAAGGATAATCTGCGCATTGCCTGGTTTGGTAATTTTCTGACAGGAGCCAGTATTTCTTTGGTTG
TACCTTTTATGCCCATCTTCGTGGAAAATCTAGGTGTAGGGAGTCAGCAAGTCGCTTTTTATGCAGGCTTAGCAAT
TTCTGTCTCTGCTATTTCCGCGGCGCTCTTTTCTCCTATTTGGGGTATTCTTGCTGACAAATACGGCCGAAAACCCA
TGATGATTCGGGCAGGTCTTGCTATGACTATCACTATGGGAGGCTTGGCCTTTGTCCCAAATATCTATTGGTTAAT
CTTTCTTCGTTTACTAAACGGTGTATTTGCAGGTTTTGTTCCTAATGCAACGGCACTGATAGCCAGTCAGGTTCCAA
AGGAGAAATCAGGCTCTGCCTTAGGTACTTTGTCTACAGGCGTAGTTGCAGGTACTCTAACTGGTCCCTTTATTGG
TGGCTTTATCGCAGAATTATTTGGCATTCGTACAGTTTTCTTACTGGTTGGTAGTTTTCTATTTTTAGCTGCTATTTT
GACTATTTGCTTTATCAAGGAAGATTTTCAACCAGTAGCCAAGGAAAAGGCTATTCCAACAAAGGAATTATTTACC
TCGGTTAAATATCCCTATCTTTTGCTCAATCTCTTTTTAACCAGTTTTGTCATCCAATTTTCAGCTCAATCGATTGG
CCCTATTTTGGCTCTTTATGTACGCGACTTAGGGCAGACAGAGAATCTTCTTTTTGTCTCTGGTTTGATTGTGTCCA
GTATGGGCTTTTCCAGCATGATGAGTGCAGGAGTCATGGGCAAGCTAGGTGACAAGGTGGGCAATCATCGTCTCTT
GGTTGTCGCCCAGTTTTATTCAGTCATCATCTATCTCCTCTGTGCCAATGCCTCTAGCCCCCTTCAACTAGGACTCT
ATCGTTTCCTCTTTGGATTGGGAACCGGTGCCTTGATTCCCGGGGTTAATGCCCTACTCAGCAAAATGACTCCCAA
AGCCGGCATTTCGAGGGTCTTTGCCTTCAATCAGGTATTCTTTTATCTGGGAGGTGTTGTTGGTCCCATGGCAGGTT
CTGCAGTAGCAGGTCAATTTGGCTACCATGCTGTCTTTTATGCGACAAGCCTTTGTGTTGCCTTTAGTTGTCTCTTT
AACCTGATTCAATTTCGAACATTATTAAAAGTAAAGGAAATCTAG

MIIMALRTSFLKCISFLKEVGKMTEINWKDNLRIAWFGNFLTGASISLVVPFMPIFVENLGVGSQQVAFYAGLAISVSAIS
AALFSPIWGILADKYGRKPMMIRAGLAMTITMGGLAFVPNIYWLIFLRLLNGVFAGFVPNATALIASQVPKEKSGSALG
TLSTGVVAGTLTGPFIGGFIAELFGIRTVFLLVGSFLFLAAILTICFIKEDFQPVAKEKAIPTKELFTSVKYPYLLLNLFLTS
FVIQFSAQSIGPILALYVRDLGQTENLLFVSGLIVSSMGFSSMMSAGVMGKLGDKVGNHRLLVVAQFYSVIIYLLCANAS
SPLQLGLYRFLFGLGTGALIPGVNALISKMTPKAGISRVFAFNQVTFYLGGVVGPMAGSAVAGQFGYHAVFYATSLCV
AFSCLFNLIQFRTLLKVKEIZ

ID36 1311bp

ATGGCCCTACCAACTATTGCCATTGTAGGACGTCCCAATGTTGGGAAATCAACCCTATTTAATCGGATCGCTGGTG
AGCGAATCTCCATTGTAGAAGATGTCGAAGGAGTGACACGTGACCGTATTTATGCAACGGGTGAGTGGCTCAATC
GTTCTTTTAGCATGATTGATACAGGAGGAATTGATGATGTCGATGCTCCTTTCATGGAACAAATCAAGCACCAGGC
AGAAATTGCCATGGAAGAAGCAGATGTTATCGTTTTTGTCGTGTCTGGTAAGGAAGGAATTACTGATGCAGACGA
ATACGTAGCTCGTAAGCTTTATAAGACCCACAAACCAGTTATCCTCGCAGTCAACAAGGTGGACAACCCTGAGAT
GAGAAATGATATATATGATTTCTATGCTCTCGGTTTGGGTGAACCATTGCCTATCTCATCTGTCCATGGAATCGGT
ACAGGGGATGTGCTAGATGCGATCGTAGAAAATCTTCCAAATGAATATGAGGAAGAAAATCCAGATGTCATTAAG
TTTAGCTTGATTGGTCGTCCTAACGTTGGAAAATCAAGCTTGATCAATGCTATCTTGGGAGAAGACCGTGTTATTG
CTAGTCCTGTTGCTGGAACAACTCGTGATGCCATTGATACCCACTTTACAGATACAGATGGTCAAGAGTTTACCAT
GATTGATACGGCTGGTATGCGTAAGTCTGGTAAGGTTTATGAAAATACTGAGAAATACTCTGTTATGCGTGCCATG
CGTGCTATTGACCGTTCAGATGTGGTCTTGATGGTCATCAATGCGGAAGAAGGCATTCGTGAGTACGACAAGCGTA
TCGCAGGATTTGCCCATGAAGCTGGTAAAGGGATGATTATCGTGGTCAACAAGTGGGATACGCTTGAAAAAGATA
ACCACACTATGAAAAACTGGGAAGAAGATATCCGTGAGCAGTTCCAATACCTGCCTTACGCACCGATTATCTTTGT
ATCAGCTTTAACCAAGCAACGTCTCCACAAACTTCCTGAGATGATTAAGCAAATCAGCGAAAGTCAAAATACACG
TATTCCATCAGCTGTCTTGAACGATGTCATCATGGATGCCATTGCCATCAACCCAACACCGACAGACAAAGGAAA
ACGTCTCAAGATTTTCTATGCGACCCAAGTGGCAACCAAACCACCAACCTTTGTCATCTTTGTCAATGAAGAAGAA
CTCATGCACTTTTCTTACCTGCGTTTCTTGGAAAATCAAATCCGCAAGGCCTTTGTTTTTGAGGGAACACCGATTCA
TCTCATCGCAAGAAAACGCAAATAA

MALPTIAIVGRPNVGKSTLFNRIAGERISIVEDVEGVTRDRIYATGEWLNRSFSMIDTGGIDDVDAPFMEQIKHQAEIAME
EADVIVFVVSGKEGITDADEYVARKLYKTHKPVILAVNKVDNPEMRNDIYDFYALGLGEPLPISSVHGIGTGDVLDAIVE
NLPNEYEEENPDVIKFSLIGRPNVGKSSLINAILGEDRVIASPVAGTTRDAIDTHFTDTDGQEFTMIDTAGMRKSGKVYE
NTEKYSVMRAMRAIDRSDVVLMVINAEEGIREYDKRIAGFAHEAGKGMIIVVNKWDTLEKDNHTMKNWEEDIREQFQ
YLPYAPIIFVSALTKQRLHKLPEMIKQISESQNTRIPSAVLNDVIMDAIAINPTPTDKGKRLKIFYATQVATKPPTFVIFVNE
EELMHFSYLRFLENQIRKAFVFEGTPIHLIARKRKZ

ID37 714bp

ATGACAGAAACCATTAAATTGATGAAGGCTCATACTTCAGTGCGCAGGTTTAAAGAGCAAGAAATTCCCCAAGTA
GACTTAAATGAGATTTTGACAGCAGCCCAGATGGCATCATCTTGGAAGAATTTCCAATCCTACTCTGTGATTGTGG
TACGAAGTCAAGAGAAGAAAGATGCCTTGTATGAATTGGTACCTCAAGAAGCCATTCGCCAGTCTGCTGTTTTCCT
TCTCTTTGTCGGAGATTTGAACCGAGCAGAAAAGGGAGCCCGACTTCATACCGACACCTTCCAACCCCAAGGTGT
GGAAGGTCTCTTGATTAGTTCGGTCGATGCAGCTCTTGCTGGACAAAACGCCTTGTTGGCAGCTGAAAGCTTGGGC
TATGGTGGTGTGATTATCGGTTTGGTTCGATACAAGTCTGAAGAAGTGGCAGAGCTCTTTAACCTACCTGACTACA
CCTATTCTGTCTTTGGGATGGCACTGGGTGTGCCAAATCAACATCATGATATGAAACCGAGACTGCCACTAGAGAA
TGTTGTCTTTGAGGAAGAATACCAAGAACAGTCAACTGAGGCAATCCAAGCTTATGACCGTGTTCAGGCTGACTAT
GCTGGGGCGCGTGCGACCACAAGCTGGAGTCAGCGCCTAGCAGAACAGTTTGGTCAAGCTGAACCAAGCTCAACT
AGAAAAAATCTTGAACAGAAGAAATTATTGTAG

MTETIKLMKAHTSVRRFKEQEIPQVDLNEILTAAQMASSWKNFQSYSVIVVRSQEKKDALYELVPQEAIRQSAVFLLFV
GDLNRAEKGARLHTDTFQPQGVEGLLISSVDAALAGQNALLAAESLGYGGVIIGLVRYKSEEVAELFNLPDYTYSVFG
MALGVPNQHHDMKPRLPLENVVFEEEYQEQSTEAIQAYDRVQADYAGARATTSWSQRLAEQFGQAEPSSTRKNLEQK
KLLZ

ID38 729bp

ATGACAGAAATTAGACTAGAGCACGTCAGTTATGCCTATGGTCAGGAGAGGATTTTAGAGGATATCAACCTACAG
GTGACTTCAGGCGAAGTGGTTTCCATCCTAGGCCCAAGTGGTGTTGGAAAGACCACCCTCTTTAATCTAATCGCTG
GGATTTTAGAAGTTCAGTCAGGGAGAATTGTCCTTGATGGTGAAGAAAATCCCAAGGGGCGCGTGAGTTATATGTT
GCAAAAGGATCTGCTCTTGGAGCACAAGACGGTGCTTGGAAATATCATTCTGCCCCTCTTGATTCAAAAGGTGGAT
AAGGCAGAAGCTATTTCCCGAGCGGATAAAATTCTTGCGACCTTCCAGCTGACAGCTGTAAGAGACAAGTATCCT
CATGAACTTAGCGGTGGGATGCGCCAGCGTGTAGCCTTACTCCGGACCTACCTTTTGGGCACAAGCTCTTTCTCT
TAGATGAGGCCTTTAGCGCCTTGGATGAGATGACAAAGATGGAACTCCACGCTTGGTATCTTGAGATTCACAAGC
AGTTGCAGCTAACAACCCTGATCATCACGCATAGTATTGAGGAGGCCCTCAATCTCAGCGACCGTATCTATATCTT
GAAAAATCGCCCTGGGCAGATTGTTTCAGAAATTAAACTAGATTGGTCTGAAGATGAGGACAAGGAAGTCCAAAA
GATTGCCTACAAACGTCAAATTTTGGCGGAATTAGGCTTAGATAAGTAG

MTEIRLEHVSYAYGQERILEDINLQVTSGEVVSILGPSGVGKTTLFNLIAGILEVQSGRIVLDGEENPKGRVSYMLQKDLL
LEHKTVLGNILPLLIQKVDKAEAISRADKILATEQLTAVRDKYPHELSGGMRQRVALLRTYLFGHKLFLLDEAFSALDE
MTKMELHAWYLEIHKQLQLTTLIITHSIEEALNLSDRIYILKNRPGQIVSEIKLDWSEDEDKEVQKIAYKRQILAELGLDK
Z

ID39 2433bp

ATGAACTATTCAAAAGCATTGAATGAATGTATCGAAAGTGCCTACATGGTTGCTGGACATTTTGGAGCTCGTTATC
TAGAGTCGTGGCACTTGTTGATTGCCATGTCTAATCACAGTTATAGTGTAGCAGGGGCAACTTTAAATGATTATCC
GTATGAGATGGACCGTTTAGAAGAGGTGGCTTTGGAACTGACTGAAACGGACTATAGCCAGGATGAAACCTTTAC
GGAATTGCCGTTCTCCCGTCGTTTGCAGGTTCTTTTTGATGAAGCAGAGTATGTAGCGTCAGTGGTCCATGCTAAG
GTACTAGGGACAGAGCACGTCCTCTATGCGATTTTGCATGATAGCAATGCCTTGGCGACTCGTATCTTGGAGAGGG
CTGGTTTTTCTTATGAAGACAAGAAAGATCAGGTCAAGATTGCTGCTCTTCGTCGAAATTTAGAAGAACGGGCAGG
CTGGACTCGTGAAGATCTCAAGGCTTTACGCCAACGCCATCGTACAGTAGCTGACAAGCAAAATTCTATGGCCAA
TATGATGGGCATGCCGCAGACTCCTAGTGGTGGTCTCGAGGATTATACGCATGATTTGACAGAGCAAGCGCGTTCT
GGCAAGTTAGAACCAGTCATCGGTCGGGACAAGGAAATCTCACGTATGATTCAAATCTTGAGCCGGAAGACTAAG
AACAACCCTGTCTTGGTTGGGGATGCTGGTGTCGGGAAAACAGCTCTGGCGCTTGGTCTTGCCCAGCGTATTGCTA
GTGGTGACGTGCCTGCGGAAATGGCTAAGATGCGCGTGTTAGAACTTGATTTGATGAATGTCGTTGCAGGGACAC
GCTTCCGTGGTGACTTTGAAGAACGCATGAATAATATCATCAAGGATATTGAAGAAGATGGCCAAGTCATCCTCTT
TATCGATGAACTCCACACCATCATGGGTTCTGGTAGCGGGATTGATTCGACTCTGGATGCGGCCAATATCTTGAAA
CCAGCCTTGGCGCGTGGAACTTTGAGAACGGTTGGTGCCACTACTCAGGAAGAATATCAAAAACATATCGAAAAA
GATGCGGCACTTTCTCGTCGTTTCGCTAAAGTGACGATTGAAGAACCAAGTGTGGCAGATAGTATGACTATTTTAC
AAGGTTTGAAGGCGACTTATGAGAAACATCACCGTGTACAAATCACAGATGAAGCGGTTGAAACAGCGGTTAAGA
TGGCTCATCGTTATTTAACCAGTCGTCACTTGCCAGACTCTGCTATCGATCTCTTGGATGAGGCGGCAGCAACAGT
GCAAAATAAGGCAAAGCATGTAAAAGCAGACGATTCAGATTTGAGTCCAGCTGACAAGGCCCTGATGGATGGCAA
GTGGAAACAGGCAGCCCAGCTAATCGCAAAAGAAGAGGAAGTACCTGTCTACAAAGACTTGGTGACAGAGTCTGA
TATTTTGACCACCTTGAGTCGCTTGTCAGGAATCCCAGTTCAAAAACTGACTCAAACGGATGCTAAGAAGTATTTA
AATCTTGAAGCAGAACTCCATAAACGGGTTATCGGTCAAGATCAAGCTGTTTCAAGCATTAGCCGTGCCATTCGCC
GCAACCAGTCAGGGATTCGCAGTCATAAGCGTCCGATTGGTTCCTTTATGTTCCTAGGGCCTACAGGTGTCGGGAA
AACTGAATTAGCCAAGGCTCTGGCAGAAGTTCTTTTTGACGACGAATCAGCCCTTATCCGCTTTGATATGAGTGAG
TATATGGAGAAATTTGCAGCTAGTCGTCTCAACGGAGCTCCTCCAGGCTATGTAGGATATGAAGAAGGTGGGGAG
TTGACAGAGAAGGTTCGCAATAAACCCTATTCCGTTCTCCTCTTTGATGAGGTAGAGAAGGCCCACCCAGATATCT
TTAATGTTCTCTTGCAGGTTCTGGATGACGGTGTCTTGACAGATAGCAAGGGACGCAAGGTCGATTTTTCAAATAC
CATTATCATTATGACATCGAATCTAGGTGCGACTGCCCTTCGTGATGATAAGACTGTTGGTTTTGGGGCTAAGGAT
ATTCGTTTTGACCAGGAAAATATGGAAAAACGCATGTTTGAAGAACTGAAAAAAGCTTATAGACCGGAATTCATC
AACCGTATTGATGAGAAGGTGGTCTTCCATAGCCTATCTAGTGATCATATGCAGGAAGTGGTGAAGATTATGGTCA
AGCCTTTAGTGGCAAGTTTGACTGAAAAAGGCATTGACTTGAAATTACAAGCTTCAGCTCTGAAATTGTTAGCAAA
TCAAGGATATGACCCAGAGATGGGAGCTCGCCCACTTCGCAGAACCCTGCAAACAGAAGTGGAGGACAAGTTGGC
AGAACTTCTTCTCAAGGGAGATTTAGTGGCAGGCAGCACACTTAAGATTGGTGTCAAAGCAGGCCAGTTAAAATT
TGATATTGCATAA

MNYSKALNECIESAYMVAGHFGARYLESWHLLIAMSNHSYSVAGATLNDYPYEMDRLEEVALELTETDYSQDETFTE
LPFSRRLQVLFDEAEYVASVVHAKVLGTEHVLYAILHDSNALATRILERAGFSYEDKKDQVKIAALRRNLEERAGWTR
EDLKALRQRHRTVADKQNSMANMMGMPQTPSGGLEDYTHDLTEQARSGKLEPVIGRDKEISRMIQILSRKTKNNPVLV
GDAGVGKTALALGLAQRIASGDVPAEMAKMRVLELDLMNVVAGTRFRGDFEERMNNIIKDIEEDGQVILFIDELHTIM
GSGSGIDSTLDAANILKPALARGTLRTVGATTQEEYQKHIEKDAALSRRFAKVTIEEPSVADSMTILQGLKATYEKHHRV
QITDEAVETAVKMAHRYLTSRHLPDSAIDLLDEAAATVQNKAKHVKADDSDLSPADKALMDGKWKQAAQLIAKEEEV
PVYKDLVTESDILTTLSRLSGIPVQKLTQTDAKKYLNLEAELHKRVIGQDQAVSSISRAIRRNQSGIRSHKRPIGSFMFLGP
TGVGKTELAKALAEVLFDDESALIRFDMSEYMEKFAASRLNGAPPGYVGYEEGGELTEKVRNKPYSVLLFDEVEKAHP
DIFNVLLQVLDDGVLTDSKGRKVDFSNTIIIMTSNLGATALRDDKTVGFGAKDIRFDQENMEKRMFEELKKAYRPEFIN
RIDEKVVFHSLSSDHMQEVVKIMVKPLVASLTEKGIDLKLQASALKLLANQGYDPEMGARPLRRTLQTEVEDKLAELLL
KGDLVAGSTLKIGVKAGQLKFDIAZ

ID40 1008bp

ATGAAGAAAACATGGAAAGTGTTTTTAACGCTTGTAACAGCTCTTGTAGCTGTTGTGCTTGTGGCCTGTGGTCAAG
GAACTGCTTCTAAAGACAACAAAGAGGCAGAACTTAAGAAGGTTGACTTTATCCTAGACTGGACACCAAATACCA
ACCACACAGGGCTTTATGTTGCCAAGGAAAAAGGTTATTTCAAAGAAGCTGGAGTGGATGTTGATTTGAAATTGC
CACCAGAAGAAAGTTCTTCTGACTTGGTTATCAACGGAAAGGCACCATTTGCAGTGTATTTCCAAGACTACATGGC
TAAGAAATTGGAAAAAGGAGCAGGAATCACTGCCGTTGCAGCTATTGTTGAACACAATACATCAGGAATCATCTC
TCGTAAATCTGATAATGTAAGCAGTCCAAAAGACTTGGTTGGTAAGAAATATGGGACATGGAATGACCCAACTGA
ACTTGCTATGTTGAAAACCTTGGTAGAATCTCAAGGTGGAGACTTTGAGAAGGTTGAAAAAGTACCAAATAACGA
CTCAAACTCAATCACACCGATTGCCAATGGCGTCTTTGATACTGCTTGGATTTACTACGGTTGGGATGGTATCCTT
GCTAAATCTCAAGGTGTAGATGCTAACTTCAATGTACTTGAAAGACTATGTCAAGGAGTTTGACTACTATTCACCAG
TTATCATCGCAAACAACGACTATCTGAAAGATAACAAAGAAGAAGCTCGCAAAGTCATCCAAGCCATCAAAAAAG
GCTACCAATATGCCATGGAACATCCAGAAGAAGCTGCAGATATTCTCATCAAGAATGCACCTGAACTCAAGGAAA
AACGTGACTTTGTCATCGAATCTCAAAAATACTTGTCAAAAGAATACGCAAGCGACAAGGAAAAATGGGGTCAAT
TTGACGCAGCTCGCTGGAATGCTTTCTACAAATGGGATAAAGAAAATGGTATCCTTAAAGAAGACTTGACAGACA
AAGGCTTCACCAACGAATTTGTGAAATAA

MKKTWKVFLTLVTALVAVVLVACGQGTASKDNKEAELKKVDFILDWTPNTNHTGLYVAKEKGYFKEAGVDVDLKLP
PEESSSDLVINGKAPFAVYFQDYMAKKLEKGAGITAVAAIVEHNTSGIISRKSDNVSSPKDLVGKKYGTWNDPTELAML
KTLVESQGGDFEKVEKVPNNDSNSITPIANGVFDTAWIYYGWDGILAKSQGVDANFMYLKDYVKEFDYYSPVIIANND
YLKDNKEEARKVIQAIKKGYQYAMEHPEEAADILIKNAPELKEKRDFVIESQKYLSKEYASDKEKWGQFDAARWNAFY
KWDKENGILKEDLTDKGFTNEFVKZ

ID41 762bp

TTGATGAGAAACTTGAGAAGTATACTGAGACGACACATTAGTCTATTGGGCTTTCTCGGAGTATTGTCAATCTGGC
AGTTAGCAGGTTTTCTTAAACTTCTCCCCAAGTTTATCCTGCCGACACCTCTTGAAATTCTCCAGCCCTTTGTTCGT
GACAGAGAATTTCTCTGGCACCATAGCTGGGCGACCTTGAGAGTGGCTTTACTGGGGCTGATTTTGGGAGTTTTGA
TTGCCTGTCTTATGGCTGTGCTCATGGATAGTTTGACTTGGCTCAATGACCTGATTTACCCTATGATGGTGGTCATT
CAGACCATTCCGACCATTGCCATAGCTCCTATCCTGGTCTTGTGGCTAGGTTATGGGATTTTGCCCAAGATTGTCTT
GATTATCTTAACGACAACCTTTCCCATCATCGTTAGTATTTTGGACGGTTTTAGGCATTGCGACAAGGATATGCTG
ACCTTGTTTAGTCTGATGCGGGCCAAGCCTTGGCAAATCCTGTGGCATTTTAAAATCCCAGTTAGCCTGCCTTACT
TTTATGCAGGTCTGAGGGTCAGTGTCTCCTACGCCTTTATCACAACTGTGGTATCTGAGTGGTTGGGAGGTTTTGA
AGGTCTTGGTGTTTATATGATTCAGTCTAAAAAACTGTTTCAGTATGATACCATGTTTGCCATTATTATTCTGGTGT
CGATTATCAGTCTTTTGGGTATGAAGCTGGTCGATATCAGTGAAAAATATGTGATTAAATGGAAACGTTCGTAG

MMRNLRSILRRHISLLGFLGVLSIWQLAGFLKLLPKFILPTPLEILQPFVRDREFLWHHSWATLRVALLGLILGVLIACLM
AVLMDSLTWLNDLIYPMMVVIQTIPTIAIAPILVLWLGYGILPKIVLIILTTTFPIIVSILDGFRHCDKDMLTLFSLMRAKP
WQILWHFKIPVSLPYFYAGLRVSVSYAFITTVVSEWLGGFEGLGVYMIQSKKLFQYDTMFAIIILVSIISLLGMKLVDISEK
YVIKWKRSZ

ID42 372bp

TTGATTTTTAATCCTATTTGCTGTATGATAAGGGAAAAGAAAGGGGACAGAGATATGGCTTTTACCAATACCCACA
TGCGATCTGCTAGTTTTGGTATTGTTACCAGCTTGCCTGATGACATCATTGACTCTTTTTGGTATATCATCGACCAT
TTCTTAAAAAATGTCTTTGAATTGGAAGAAGAACTCGAGTTTCAATTGCTTAATAACCAAGGAAAGATTACCTTCC
ACTTTTCAAGTCAACACCTCCCTACAGCCATTGATTTTGACTTTAACCATCCTTTCGACCCTCGTTATCCCCCAAGA
GTACTGGTTTTAGACATGGACGGTAGAGAAACTATCCTCCTCCCAGAAGAAAATGACCTATTTTAA

MIFNPICCMIREKKGDRDMAFTNTHMRSASFGIVTSLPDDIIDSFWYIIDHFLKNVFELEEELEFQLLNNQGKITFHFSSQ
HLPTAIDFDFNHPFDPRYPPRVLVLDMDGRETILLPEENDLFZ

ID43 1569bp

ACAGCGGTGTCATTCTATCTATTTTAAGAAAAGTAATAATCAATTGTTAAAAATAGTAAAAAAATTGGAGGTTCTG
ATGAAATATTTTGTTCCTAATGAGGTATTCAGTATTCGTAAATTAAAGGTGGGGACTTGCTCGGTACTATTGGCAA
TTTCAATTTTGGGAAGCCAAGGTATTTTATCGGATGAAGTTGTTACTAGTTCTTCACCGATGGCTACAAAAGAGTC
TTCTAATGCAATTACTAATGATTTAGATAATTCACCAACTGTTAATCAGAATCGTTCTGCTGAAATGATTGCCTCTA
ATTCAACCACTAATGGTTTAGATAATTCGTTAAGTGTTAATAGCATCAGCTCTAATGGTACTATTCGTTCCAATTCA
CAATTAGACAACAGAACAGTTGAATCTACAGTAACATCTACTAATGAAAATAAGAGTTATAAGGAAGATGTTATA
AGTGACAGAATTATCAAAAAAGAATTTGAAGATACTGCTTTAAGTGTAAAAGATTATGGTGCAGTAGGTGATGGG
ATTCATGATGATCGACAAGCAATTCAAGATGCAATAGATGCTGCAGCTCAAGGGCTAGGTGGAGGAAATGTATAT
TTTCCTGAAGGAACTTATTTAGTAAAAGAAATTGTTTTTTTAAAAAGTCATACACACTTAGAATTGAATGAGAAAG
CTACAATTCTAAATGGTATAAATATTAAGAATCACCCTTCCATTGTTTTTATGACAGGTTTATTTACGGATGATGGT
GCGCAAGTAGAATGGGGCCCAACAGAAGATATTAGTTATTCTGGTGGTACGATTGATATGAACGGTGCTTTGAAT
GAAGAAGGAACTAAAGCAAAAAATCTACCACTTATAAATTCTTCAGGTGCATTTGCTATTGGGAATTCAAATAAC
GTAACTATAAAAAATGTAACATTCAAGGATAGTTATCAAGGGCATGCTATTCAAATTGCAGGTTCGAAAAATGTAT
TAGTTGATAATTGTCGTTTTCTTGGGCAAGCCTTACCCAAAACGATGAAGGATGGGCAAATCATAAGTAAGGAGA
GCATTCAGATTGAACCATTAACTAGAAAAGGTTTTCCTTATGCCTTGAATGATGATGGGAAAAAATCTGAAAATGT
GACTATTCAAAATTCCTATTTTGGCAAAAGTGATAAATCTGGGGAATTAGTAACAGCAATTGGCACACACTATCAA
ACATTGTCGACACAGAACCCCTCTAATATTAAAATTGAAAATAATCATTTTGATAACATGATGTATGCAGGTGTAC
GTTTTACAGGATTCACTGATGTATTAATCAAAGGAAATCGCTTTGATAAGAAAGTTAAAGGAGAGAGTGTACATTA
TCGAGAAAGCGGAGCAGCTTTAGTAAATGCTTATAGCTATAAAAACACTAAAGACCTATTAGATTTAAATAAACA
GGTGGTTATCGCCGAAAATATAATTTAATATTGCCGATCCTAAAACAAAAGCGATACGAGTTGCAAAAGATAGTGC
AGAATGTTTAGGAAAAGTATCAGATATTACTGTAACAAAAAATGTAATTAATAATAATTCTAAGGAAACAGAACA
ACCAAATATTGAATTATTACGAGTTAGTGATAATTTAGTAGTCTCAGAGAATAGT

QRCHSIYFKKSNNQLLKIVKKLEVLMKYFVPNEVFSIRKLKVGTCSVLLAISILGSQGILSDEVVTSSSPMATKESSNAITN
DLDNSPTVNQNRSAEMIASNSTTNGLDNSLSVNSISSNGTIRSNSQLDNRTVESTVTSTNENKSYKEDVISDRIIKKEFEDT
ALSVKDYGAVGDGIHDDRQAIQDAIDAAAQGLGGGNVYFPEGTYLVKEIVFLKSHTHLELNEKATILNGINIKNHPSIVF
MTGLFTDDGAQVEWGPTEDISYSGGTIDMNGALNEEGTKAKNLPLINSSGAFAIGNSNNVTIKNVTFKDSYQGHAIQIA

GSKNVLVDNSRFLGQALPKTMKDGQIISKESIQIEPLTRKGFPYALNDDGKKSENVTIQNSYFGKSDKSGELVTAIGTHY
QTLSTQNPSNIKIQNNHFDNMMYAGVRFTGFTDVLIKGNRFDKKVKGESVHYRESGAALVNAYSYKNTKDLLDLNKQ
VVIAENIFNIADPKTKAIRVAKDSAECLGKVSDITVTKNVINNNSKETEQPNIELLRVSDNLVVSENS

__ID44 324bp__

GTGATGAAAGAAACTCAGCTATTAAAAGGTGTTCTTGAAGGTTGTGTCTTGGATATGATTGGTCAAAAAGAGCGGT
ATGGTTATGAGTTGGTTCAGACTTTGCGAGAGGCTGGATTTGATACTATCGTTCCAGGAACTATTTATCCTTTGTTG
CAAAAGTTAGAAAAAAATCAATGGATAAGAGGCGACATGCGCCCGTCGCCAGATGGTCCAGATCGGAAGTATTTT
TCATTAATGAAAGAAGGAGAAGAGCGTGTCTCAGTCTTTTGGCAACAATGGGACGATTTGAGTCAAAAAGTAGAA
GGGATTAAGAATGGGGGTTAA

MMKETQLLKGVLEGCVLDMIGQKERYGYELVQTLREAGFDTIVPGTIYPLLQKLEKNQWIRGDMRPSPDGPDRKYFSL
MKEGEERVSVFWQQWDDLSQKVEGIKNGGZ

__ID45 816bp__

ATGAAGAAAATGAAGTATTACGAAGAAACAAGCGCTTTGCTACATGAGTTTTCTGAGGAGAATCAAAAGTATTTT
GAGGAGTTGTGGGAAAGTTTTAATCTTGCTGGATTTCTCTATGATGAAGACTATCTCAGAGAGCAGATCTATTTGA
TGATGCTAGATTTCTCAGAAGCAGAACGAGATGGCATGAGTGCAGAGGATTATCTAGGTAAGAATCCTAAAAAAA
TAATGAAAGAGATTCTCAAGGGAGCACCTCGCAGTTCTATCAAAGAGTCCCTTTTGACGCCAATTCTTGTCCTGGC
GGTATTACGTTATTATCAACTACTAAGTGATTTTTCTAAAGGTCCTCTCTTAACAGTCAATTTGCTCACATTTTTAG
GGCAACTTCTTATTTTTCTGATTGGATTTGGACTTGTGGCCACAATTTTACGAAGAAGTTTAGTCCAAGATTCTCCT
AAAATGAAAATTGGCACTTACATTGTTGTTGGGACTATAGTTCTTCTAGTTGTTTTAGGATATGTAGGAATGGCAA
GCTTCATACAAGAAGGAGCCTTTTATATTCCGGCTCCCTGGGATAGTTTGTCTGTCTTTACGATTTCGCTAGTTATC
GGTATTTGGAATTGGAAAGAAGCGGTCTTTCGTCCATTTGTCAGTATGATTATTGCCCATCTTGTGGTGGGTTCTCT
GCTCCGTTATTATGAGTGGATGGGAATTTCAAATGTTTTCCTTACAAAAGTTATTCCTTTAGCTGTCCTCTTTATTG
GAATCTTTGTCTTGTTCCGTGGGTTTAAGAAGATAAAATGGAGTGAAGTATAG

MKKMKYYEETSALLHEFSEENQKYFEELWESFNLAGFLYDEDYLREQIYLMMLDFSEAERDGMSAEDYLGKNPKKIM
KEILKGAPRSSIKESLLTPILVLAVLRYYQLLSDFSKGPLLTVNLLTFLGQLLIFLIGFGLVATILRRSLVQDSPKMKIGTYI
VVGTIVLLVVLGYVGMASFIQEGAFYIPAPWDSLSVFTISLVIGIWNWKEAVFRPFVSMIIAHLVVGSLLRYYEWMGISN
VFLTKVIPLAVLFIGIFVLFRGFKKIKWSEVZ

__ID46 348bp__

CTGTTTTTTTATTTATACTCAATGAAAATCAAAGAGCAAACTAGGAAGCTAGCCGCAGGTTGCTCAAAACACTGTT
TTGAGGTTGTAGACGAAACTGACGAAGTCAGCTCAAAACATGTTTTTGAGGTTGTAGATGAAACTGACGAAGTCA
GCTCAAAACACTGTTTTGAGGTTGTAGATGAAACTGACGAAGTCAGCTCAAAACACTGTTTTGAGGTTGTAGATGA
AACTGACGAAGTCAGCTCAAAACATGTTTTTGAGGTTGTAGATGAAACTGACGAAGTCAGTAACCATACATACGG
TAGGGCGACGCTGACGTGGTTTGAAGAGATTTTCGAAGAGTATTAA

MFFYLYSMKIKEQTRKLAAGCSKHCFEVVDETDEVSSKHVFEVVDETDEVSSKHCFEVVDETDEVSSKHCFEVVDETD
EVSSKHVFEVVDETDEVSNHTYGRATLTWFEEIFEEYZ

__ID47 1260bp__

ATGCAGAATCTGAAATTTGCCTTTTCATCTATCATGGCTCACAAGATGCGTTCTTTGCTTACTATGATTGGGATTAT
TATCGGTGTTTCATCAGTTGTTGTGATTATGGCTTTGGGTGATTCCCTATCTCGTCAAGTCAATAAAGATATGACTA
AATCTCAGAAAAATATTAGCGTCTTTTTCTCTCCTAAAAAAAGTAAAGACGGGTCTTTTACTCAGAAACAATCAGC
TTTTACGGTTTCTGGAAAGGAAGAGGAAGTTCCTGTTGAACCGCCAAAACCGCAAGAATCCTGGGTCCAAGAGGC
AGCTAAACTGAAGGGAGTGGATAGTTACTATGTAACCAATTCAACGAATGCCATCTTGACCTATCAAGATAAAAA
GGTTGAGAATGCTAATTTGACAGGTGGAAACAGAACTTACATGGACGCTGTTAAGAATGAAATTATTGCAGGTCG
TAGTCTGAGAGAGCAAGATTTCAAAGAGTTTGCAAGTGTCATTTTGCTAGATGAGGAATTGTCCATTAGTTTATTT
GAATCTCCTCAAGAGGCTATTAACAAGGTTGTAGAAGTCAATGGATTTAGTTACCGGGTCATTGGGGTTTATACTA
GTCCGGAGGCTAAAAGATCAAAAATATATGGGTTTGGTGGCTTGCCTATTACTACCAATATCTCCCTTGCTGCGAA
TTTTAATGTAGATGAAATAGCTAATATTGTCTTTCGAGTGAATGATACCAGTTTAACCCCAACTCTGGGTCCAGAA
CTGGCACGAAAAATGACAGAGCTTGCAGGCTTACAACAGGGAGAATACCAGGTGGCAGATGAGTCCGTTGTATTT
GCAGAAATTCAACAATCGTTTAGTTTTATGACGACGATTATTAGTTCCATCGCAGGGATTTCTCTCTTTGTTGGAG
GAACTGGTGTCATGAACATCATGCTGGTTTCGGTGACAGAGCGCACTCGTGAGATTGGTCTTCGTAAGGCTTTGGG
TGCAACACGTGCCAATATTTTAATTCAGTTTTTGATTGAATCCATGATTTTGACCTTGTTAGGTGGCTTAATTGGCT
TGACAATTGCAAGTGGTTTAACTGCCTTAGCAGGTTTGTTACTGCAAGGTTTAATAGAAGGTATAGAAGTTGGAGT
ATCAATCCCAGTCGCCCTATTTAGTCTTGCAGTTTCGGCTAGTGTTGGTATGATTTTTGGAGTCTTGCCAGCCAACA
AGGCATCGAAACTTGATCCAATTGAAGCCCTTCGTTATGAATGA

MQNLKFAFSSIMAHKMRSLLTMIGIIIGVSSVVVIMALGDSLSRQVNKDMTKSQKNISVFFSPKKSKDGSFTQKQSAFTVS
GKEEEVPVEPPKPQESWVQEAAKLKGVDSYYVTNSTNAILTYQDKKVENANLTGGNRTYMDAVKNEIIAGRSLREQDF
KEFASVILLDEELSISLFESPQEAINKVVEVNGFSYRVIGVYTSPEAKRSKIYGFGGLPITTNISLAANFNVDEIANIVFRVN
DTSLTPTLGPELARKMTELAGLQQGEYQVADESVVFAEIQQSFSFMTTIISSIAGISLFVGGTGVMNIMLVSVTERTREIG
LRKALGATRANILIQFLIESMILTLLGGLIGLTIASGLTALAGLLLQGLIEGIEVGVSIPVALFSLAVSASVGMIFGVLPANK
ASKLDPIEALRYEZ

ID48 705bp

CTGATGAAGCAACTAATTAGTCTAAAAAATATCTTCAGAAGTTACCGTAATGGTGACCAAGAACTGCAGGTTCTCA
AAAATATCAATCTAGAAGTGAATGAGGGTGAATTTGTAGCCATCATGGGACCATCTGGGTCTGGTAAGTCCACTCT
GATGAATACGATTGGCATGTTGGATACACCAACCAGTGGAGAATATTATCTTGAAGGTCAAGAAGTGGCTGGGCT
TGGTGAAAAACAACTAGCTAAGGTCCGTAACCAACAAATCGGTTTTGTCTTTCAGCAGTTCTTTCTTCTATCGAAG
CTCAATGCTCTGCAAAATGTAGAATTGCCCTTGATTTACGCAGGAGTTTCGTCTTCAAAACGTCGCAAGTTGGCTG
AGGAATATTTAGACAAGGTTGAATTGACAGAACGTAGTCACCATTTACCTTCAGAATTATCTGGTGGTCAAAAGCA
ACGTGTAGCCATTGCGCGTGCCTTGGTAAACAATCCTTCTATTATCCTAGCGGATGAACCGACAGGAGCCTTGGAT
ACCAAAACAGGTAACCAAATTATGCAATTATTGGTTGATTTGAATAAAGAAGGAAAAACCATTATCATGGTAACG
CATGAGCCTGAGATTGCTGCCTATGCCAAACGTCAGATTGTCATTCGGGATGGGGTCATTTCGTCTGACAGTGCTC
AGTTAGGAAAGGAGGAAAACTAA

MMKQLISLKNIFRSYRNGDQELQVLKNINLEVNEGEFVAIMGPSGSGKSTLMNTIGMLDTPTSGEYYLEGQEVAGLGEK
QLAKVRNQQIGFVFQQFFLLSKLNALQNVELPLIYAGVSSSKRRKLAEEYLDKVELTERSHHLPSELSGGQKQRVAIARA
LVNNPSIILADEPTGALDTKTGNQIMQLLVDLNKEGKTIIMVTHEPEIAAYAKRQIVIRDGVISSDSAQLGKEENZ

ID49 1200bp

ATGAAGAAAAAGAATGGTAAAGCTAAAAAGTGGCAACTGTATGCAGCAATCGGTGCTGCGAGTGTAGTTGTATTG
GGTGCTGGGGGGATTTTACTCTTTAGACAACCTTCTCAGACTGCTCTAAAAGATGAGCCTACTCATCTTGTTGTTG
CCAAGGAAGGAAGCGTGGCCTCCTCTGTTTTATTGTCAGGGACAGTAACAGCAAAAAATGAACAATATGTTTATTT
TGATGCTAGTAAGGGTGATTTAGATGAAATCCTTGTTTCTGTGGGCGATAAGGTCAGCGAAGGGCAGGCTTTAGTC
AAGTACAGTAGTTCAGAAGCGCAGGCGGCCTATGATTCAGCTAGTCGAGCAGTAGCTAGGGCAGATCGTCATATC
AATGAACTCAATCAAGCACGAAATGAAGCCGCTTCAGCTCCGGCTCCACAGTTACCAGCGCCAGTAGGAGGAGAA
GATGCAACGGTGCAAAGCCCAACTCCAGTGGCTGGAAATTCTGTTGCTTCTATTGACGCTCAATTGGGTGATGCCC
GTGATGCGCGTGCAGATGCTGCGGCGCAATTAAGCAAGGCTCAAAGTCAATTGGATGCAACAACTGTTCTCAGTA
CCCTAGAGGGAACTGTGGTCGAAGTCAATAGCAATGTTTCTAAATCTCCAACAGGGGCGAGTCAAGTTATGGTTC
ATATTGTCAGCAATGAAAATTTACAAGTCAAGGGAGAATTGTCTGAGTACAATCTAGCCAACCTTTCTGTAGGTCA
AGAAGTAAGCTTTACTTCTAAAGTGTATCCTGATAAAAAATGGACTGGGAAATTAAGCTATATTTCTGACTATCCT
AAAAACAATGGTGAAGCAGCTAGTCCAGCAGCCGGGAATAATACAGGTTCTAAATACCCTTATACTATTGATGTG
ACAGGCGAGGTTGGTGATTTGAAACAAGGTTTTTCTGTCAACATTGAGGTTAAAAGCAAAACTAAGGCTATTCTTG
TTCCTGTTAGCAGTCTAGTAATGGATGATAGTAAAAATTATGTCTGGATTGTGGATGAACAACAAAAGGCTAAAA
AAGTTGAGGTTTCATTGGGAAATGCTGACGCAGAAAATCAAGAAATCACTTCTGGTTTAACGAACGGTGCTAAGG
TCATCAGTAATCCAACATCTTCCTTGGAAGAAGGAAAAGAGGTGAAGGCTGATGAAGCAACTAATTAG

MKKKNGKAKKWQLYAAIGAASVVVLGAGGILLFRQPSQTALKDEPTHLVVAKEGSVASSVLLSGTVTAKNEQYVYFD
ASKGDLDEILVSVGDKVSEGQALVKYSSSEAQAAYDSASRAVARADRHINELNQARNEAASAPAPQLPAPVGGEDATV
QSPTPVAGNSVASIDAQLGDARDARADAAAQLSKAQSQLDATTVLSTLEGTVVEVNSNVSKSPTGASQVMVHIVSNEN
LQVKGELSEYNLANLSVGQEVSFTSKVYPDKKWTGKLSYISDYPKNNGEAASPAAGNNTGSKYPYTIDVTGEVGDLKQ
GFSVNIEVKSKTKAILVPVSSLVMDDSKNYVWIVDEQQKAKKVEVSLGNADAENQEITSGLTNGAKVISNPTSSLEEGKE
VKADEATNZ

ID50 759bp

ATGTCACGTAAACCATTTATCGCTGGTAACTGGAAAATGAACAAAAATCGAGAAGAAGCTAAAGCATTCGTTGAA
GCAGTTGCATCAAAACTTCCTTCATCAGATCTTGTTGAAGCAGGTATCGCTGCTCCAGCTCTTGATTTGACAACTG
TTCTTGCTGTTGCAAAAGGCTCAAACCTTAAAGTTGCTGCTCAAAACTGCTACTTTGAAAATGCAGGTGCTTTCAC
TGGTGAAACTAGCCCACAAGTTTTGAAAGAAATCGGTACTGACTACGTTGTTATCGGTCACTCAGAACGCCGTGAC
TACTTCCATGAAACTGATGAAGATATCAACAAAAAAGCAAAAGCAATCTTTGCGAACGGTATGCTTCCAATCATC
TGTTGTGGTGAATCACTTGAAACTTACGAAGCTGGTAAAGCTGCTGAATTCGTAGGTGCTCAAGTATCTGCTGCAT
TGGCTGGATTGACTGCTGAACAAGTTGCTGCCTCAGTTATCGCTTATGAGCCAATCTGGGCTATCGGTACTGGTAA
ATCAGCTTCACAAGACGATGCACAAAAAATGTGTAAAGTTGTTCGTGACGTTGTAGCTGCTGACTTTGGTCAAGAA
GTCGCAGACAAAGTTCGTGTTCAATACGGTGGTTCTGTTAAACCTGAAAATGTTGCTTCATACATGGCTTGCCCAG
ACGTTGACGGTGCCCTTGTAGGTGGTGCGTCACTTGAAGCTGAAAGCTTCTTGGCTTTGCTTGACTTTGTAAAATA
A

MSRKPFIAGNWKMNKNPEEAKAFVEAVASKLPSSDLVEAGIAAPALDLTTVLAVAKGSNLKVAAQNCYFENAGAFTG
ETSPQVLKEIGTDYVVIGHSERRDYFHETDEDINKKAKAIFANGMLPIICCGESLETYEAGKAAEFVGAQVSAALAGLTA
EQVAASVIAYEPIWAIGTGKSASQDDAQKMCKVVRDVVAADFGQEVADKVRVQYGGSVKPENVASYMACPDVDGAL
VGGASLEAESFLALLDFVKZ

ID51 1473bp

TTGAAAACAAAAATTGGATTAGCAAGTATCTGTTTACTAGGCTTGGCAACTAGTCATGTCGCTGCAAATGAAACTG
AAGTAGCAAAAACTTCGCAGGATACAACGACAGCTTCAAGTAGTTCAGAGCAAAATCAGTCTTCTAATAAAACGC
AAACGAGCGCAGAAGTACAGACTAATGCTGCTGCCCACTGGGATGGGGATTATTATGTAAAGGATGATGGTTCTA
AAGCTCAAAGTGAATGGATTTTTGACAACTACTATAAGGCTTGGTTTTATATTAATTCAGATGGTCGTTACTCGCA
GAATGAATGGCATGGAAATTACTACCTGAAATCAGGTGGATATATGGCCCAAAACGAGTGGATCTATGACAGTAA
TTACAAGAGTTGGTTTTATCTCAAGTCAGATGGGGCTTATGCTCATCAAGAATGGCAATTGATTGGAAATAAGTGG
TACTACTTCAAGAAGTGGGGTTACATGGCTAAAAGCCAATGGCAAGGAAGTTATTTCTTGAATGGTCAAGGAGCT
ATGATGCAAAATGAATGGCTCTATGATCCAGCCTATTCTGCTTATTTTTATCTAAAATCCGATGGAACTTATGCTA
ACCAAGAGTGGCAAAAAGTGGGCGGCAAATGGTACTATTTCAAGAAGTGGGGCTATATGGCTCGGAATGAGTGGC
AAGGCAACTACTATTTGACTGGAAGTGGTGCCATGGCGACTGACGAAGTGATTATGGATGGTACTCGCTATATCTT
TGCGGCCTCTGGTGAGCTCAAAGAAAAAAAAGATTTGAATGTCGGCTGGGTTCACAGAGATGGTAAGCGCTATTT
CTTTAATAATAGAGAAGAACAAGTGGGAACCGAACATGCTAAGAAAGTCATTGATATTAGTGAGCACAATGGTCG
TATCAATGATTGGAAAAAGGTTATTGATGAGAACGAAGTGGATGGTGTCATTGTTCGTCTAGGTTATAGCGGTAAA
GAAGACAAGGAATTGGCGCATAACATTAAGGAGTTAAACCGTCTGGGAATTCCTTATGGTGTCTATCTCTATACCT
ATGCTGAAAATGAGACCGTGCTGAGAGTGACGCTAAACAGACCATTGAACTTATAAAGAAATACAATATGAACCT
GTCTTACCCTATCTATTATGATGTTGAGAATTGGGAATATGTAAATAAGAGCAAGAGAGCTCCAAGTGATACAGG
CACTTGGGTTAAAATCATCAACAAGTACATGGACACGATGAAGCAGGCGGGTTATCAAAATGTGTATGTCTATAG
CTATCGTAGTTTATTACAGACGCGTTTAAAACACCCAGATATTTTAAAACATGTAAACTGGGTAGCGGCCTATACG
AATGCTTTAGAATGGGAAAACCCTCATTATTCAGGAAAAAAAGGTTGGCAATATACCTCTTCTGAATACATGAAA
GGAATCCAAGGGCGCGTAGATGTCAGCGTTTGGTATTAA

MKTKIGLASICLLGLATSHVAANETEVAKTSQDTTTASSSSEQNQSSNKTQTSAEVQTNAAAHWDGDYYVKDDGSKAQ
SEWIFDNYYKAWFYINSDGRYSQNEWHGNYYLKSGGYMAQNEWIYDSNYKSWFYLKSDGAYAHQEWQLIGNKWYY
FKKWGYMAKSQWQGSYFLNGQGAMMQNEWLYDPAYSAYFYLKSDGTYANQEWQKVGGKWYYFKKWGYMARNE
WQGNYYLTGSGAMATDEVIMDGTRYIFAASGELKEKKDLNVGWVHRDGKRYFFNNREEQVGTEHAKKVIDISEHNGR
INDWKKVIDENEVDGVIVRLGYSGKEDKELAHNIKELNRLGIPYGVYLYTYAENETDAESDAKQTIELIKKYNMNLSYPI
YYDVENWEYVNKSKRAPSDTGTWVKIINKYMDTMKQAGYQNVYVYSYRSLLQTRLKHPDILKHVNWVAAYTNALE
WENPHYSGKKGWQYTSSEYMKGIQGRVDVSVWYZ

ID52 774bp

ATGAAAAAATTTGCCAACCTTTATCTGGGACTGGTCTTTCTGGTCCTCTACCTGCCTATCTTTTACTTGATTGGCTA
TGCCTTTAATGCTGGTGATGATATGAATAGCTTTACAGGTTTTAGCTGGACTCACTTTGAAACCATGTTTGGAGAT
GGGAGACTCATGCTGATTTTGGCTCAGACATTTTTCTTGGCCTTCCTATCAGCCTTGATAGCGACCATTATCGGGA
CTTTTGGTGCCATTTACATCTACCAGTCTCGTAAGAAATACCAAGAAGCCTTTCTATCACTCAATAATATCCTCAT
GGTTGCGCCTGACGTTATGATTGGTGCTAGCTTCTTGATTCTCTTTACCCAACTCAAGTTTTCACTTGGCTTTTTGA
CCGTTCTATCTAGTCACGTGGCCTTCTCCATTCCTATCGTGGTCTTGATGGTCTTGCCTCGACTCAAGGAAATGAAT
GGCGACATGATTCATGCGGCCTATGACTTGGGAGCTAGTCAATTTCAGATGTTCAAGGAAATCATGCTTCCTTACC
TGACTCCGTCTATCATTACTGGTTATTTCATGGCCTTCACCTATTCGTTAGATGACTTTGCCGTGACCTTCTTTGTA
ACAGGAAATGGCTTTTCAACCCTATCAGTCGAGATTTACTCTCGTGCTCGCAAGGGGATTTCCTTAGAAATCAATG
CCCTGTCTGCTCTAGTCTTTCTCTTTAGTATTATCCTAGTTGTAGGTTATTACTTTATCTCTCGTGAGAAGGAGGAG
CAAGCATGA

MKKFANLYLGLVFLVLYLPIFYLIGYAFNAGDDMNSFTGFSWTHFETMFGDGRLMLILAQTFFLAFLSALIATIIGTFGA
IYIYQSRKKYQEAFLSLNNILMVAPDVMIGASFLILFTQLKFSLGFLTVLSSHVAFSIPIVVLMVLPRLKEMNGDMIHAAY
DLGASQFQMFKEIMLPYLTPSIITGYFMAFTYSLDDFAVTFFVTGNGFSTLSVEIYSRARKGISLEINALSALVFLFSIILVV
GYYFISREKEEQAZ

ID59 1071bp

ATGAAAAAAATCTATTCATTTTTAGCAGGAATTGCAGCGATTATCCTTGTCTTGTGGGGAATTGCGACTCATTTAG
ATACTAAAATCAATAGTCGACATAGTCAAAAATTGGTTATCTATAACTGGGGAGACTATATCGATCCTGAACTCTT
GACTCAGTTTACAGAAGAAACAGGAATTCAAGTTCAGTACGAGACTTTTGACTCCAACGAAGCCATGTACACTAA
GATAAAGCAGGGTGGAACGACCTACGATATTGCCATTCCAAGTGAATACGATCATTAACAAGATGAAGGACGAAGA
CCTCTTGGTTCCGCTTCATTATTCAAAAATTGAAGGAATCGAAAATATCGGACCAGAGTTTCTCAACCAGTCCTTT
GACCCAGGTAATAAATTCTCCATCCCTTACTTCTGGGGAACCTTAGGAATTGTCTACAACGAAACCATGGTAGATG

AAGCGCCTGAGCATTGGGATGACCTTTGGAAGCCGGAGTATAAGAATTCTATCATGCTCTTTGATGGGGCGCGTGA
GGTGCTGGGACTAGGACTCAATTCCCTCGGCTACAGCCTCAACTCCAAGGATCTGCAGCAGTTGGAAGAGACAGT
GGATAAGCTCTACAAACTGACTCCAAATATCAAGGCTATCGTTGCGGACGAGATGAAGGGCTATATGATTCAGAA
TAATGTTGCAATCGGCGTGACCTTCTCTGGTGAAGCCAGCCAAATGTTAGAAAAAAATGAAAATCTACGTTATGTG
GTACCGACAGAGGCCAGCAATCTTTGGTTTGACAATATGGTCATTCCCAAAACAGTTAAAAACCAAAACTCAGCC
TATGCCTTTATCAACTTTATGTTGAAACCTGAAAATGCTCTCCAAAATGCGGAGTATGTCGGCTATTCAACACCAA
ACCTACCAGCGAAGGAATTGCTCCCAGAGGAAACAAAGGAAGATAAGGCCTTCTATCCCGATGTTGAAACCATGA
AACACCTAGAAGTTTATGAGAAATTTGACCATAAATGGACAGGGAAATATAGCGACCTCTTCCTACAGTTTAAAA
TGTATCGGAAGTAG

MKKIYSFLAGIAAIILVLWGIATHLDSKINSRDSQKLVIYNWGDYIDPELLTQFTEETGIQVQYETFDSNEAMYTKIKQGG
TTYDIAIPSEYMINKMKDEDLLVPLDYSKIEGIENIGPEFLNQSFDPGNKFSIPYFWGTLGIVYNETMVDEAPEHWDDLW
KPEYKNSIMLFDGAREVLGLGLNSLGYSLNSKDLQQLEETVDKLYKLTPNIKAIVADEMKGYMIQNNVAIGVTFSGEAS
QMLEKNENLRYVVPTEASNLWFDNMVIPKTVKNQNSAYAFINFMLKPENALQNAEYVGYSTPNLPAKELLPEETKED
KAFYPDVETMKHLEVYEKFDHKWTGKYSDLFLQFKMYRKZ

ID61 1851bp

ATGAATAAAAAACTAACAGATTATGTGATTGATCTGGTGGAAATTTTAAATAAACAACAAAAGCAGGTTTTCTGG
GGAATATTTGATATTTTCAGTATGGTGGTTTCCATCATTGTATCTTATATTTTATTTTATGGGCTGATTAATCCAGC
ACCTGTTGACTACATTATCTATACGAGTTTGGCCTTCCTGTTCTATCAATTGATGATTGGTTTTTGGGGGTTGAACG
CGAGCATTAGTCGTTACAGCAAGATTACGGATTTCATGAAAATCTTTTTTGGTGTGACTGCTAGCAGTGTCTTGTC
ATATAGTATCTGTTATGCCTTCTTGCCACTCTTCTCCATCCGTTTCATCATTCTCTTTATCTTGTTGAGTACCTTCTT
GATTTTATTGCCACGGATTACTTGGCAGTTAATCTACTCCAGACGCAAAAAAGGTAGTGGTGATGGAGAACACCGT
CGGACCTTCTTGATTGGTGCCGGTGATGGTGGGGCTCTTTTTATGGATAGTTACCAACATCCAACCAGTGAATTAG
AACTGGTCGGTATTTTGGATAAGGATTCTAAGAAAAAGGGTCAAAAACTTGGTGGTATTCCTGTTTTGGGCTCTTA
TGACAATCTGCCTGAATTAGCCAAACGCCATCAAATCGAGCGTGTCATCGTTGCGATTCCGTCGCTGGATCCGTCA
GAATATGAGCGTATCTTGCAGATGTGTAATAAGCTGGGTGTCAAATGTTACAAGATGCCTAAGGTTGAAACTGTTG
TTCAGGGCCTTCACCAAGCAGGTACTGGCTTCCAAAAAATTGATATTACGGACCTTTTGGGTCGTCAGGAAATCCG
TCTTGACGAATCGCGTCTGGGTGCAGAACTGACAGGTAAGACCATCTTAGTCACAGGAGCTGGAGGTTCAATCGG
TTCTGAAATCTGTCGTCAAGTTAGTCGCTTCAATCCTGAACGCATTGTCTTGCTCGGTCATGGGGAAAACTCAATC
TACCTTGTTTATCATGAATTGATTCGTAAGTTCCAAGGGATTGATTATGTACCTGTGATTGCGGACATTCAAGACT
ATGATCGTTTGTTGCAAGTCTTTGAGCAGTACAAACCTGCTATTGTTTATCATGCGGCAGCCCACAAGCATGTTCC
TATGATGGAGCGCAATCCAAAAGAAGCCTTCAAAAACAATATCCGTGGAACTTACAATGTTGCTAAGGCTGTTGA
TGAAGCTAAAGTGTCTAAGATGGTTATGATTTCGACAGATAAGGCAGTCAATCCACCAAATGTTATGGGAGCAAC
CAAGCGCGTGGCGGAGTTGATTGTCACTGGCTTTAACCAACGTAGCCAATCAACCTACTGTGCAGTTCGTTTTGGG
AATGTTCTTGGTAGCCGTGGTAGTGTCATTCCAGTCTTTGAACGTCAGATTGCTGAAGGTGGGCCTGTAACGGTGA
CAGACTTCCGTATGACCCGTTACTTTATGACCATTCCAGAAGCTAGCCGTCTGGTTATCCATGCTGGTGCTTATGC
CAAAGATGGGGAAGTCTTTATCCTTGATATGGGCAAACCAGTCAAGATTTATGACTTGGCCAAGAAGATGGTGCTT
CTAAGTGGCCACACTGAAAGTGAAATTCCAATCGTTGAAGTTGGAATCCGCCCAGGTGAAAAACTCTACGAAGAA
CTCTTGGTATCAACCGAACTCGTTGATAATCAAGTTATGGATAAGATTTTCGTTGGTAAGGTTAATGTCATGCCTTT
AGAATCCATCAATCAAAAGATTGGAGAGTTCCGCACTCTCAGTGGAGATGAGTTGAAGCAAGCTATTATCGCCTTT
GCTAATCAAACAACCCACATTGAATAA

MNKKLTDYVIDLVEILNKQQKQVFWGIFDIFSMVVSIIVSYILFYGLINPAPVDYIIYTSLAFLFYQLMIGFWGLNASISRY
SKITDFMKIFFGVTASSVLSYSICYAFLPLFSIRFIILFILLSTFLILLPRITWQLIYSRRKKGSGDGEHRRTFLIGAGDGGALF
MDSYQHPTSELELVGILDKDSKKKGQKLGGIPVLGSYDNLPELAKRHQIERVIVAIPSLDPSEYERILQMCNKLGVKCYK
MPKVETVVQGLHQAGTGFQKIDITDLLGRQEIRLDESRLGAELTGKTILVTGAGGSIGSEICRQVSRFNPERIVLLGHGEN
SIYLVYHELIRKFQGIDYVPVIADIQDYDRLLQVFEQYKPAIVYHAAAHKHVPMMERNPKEAFKNNIRGTYNVAKAVDE
AKVSKMVMISTDKAVNPPNVMGATKRVAELIVTGFNQRSQSTYCAVRFGNVLGSRGSVIPVFERQIAEGGPVTVTDFR
MTRYFMTIPEASRLVIHAGAYAKDGEVFILDMGKPVKIYDLAKKMVLLSGHTESEIPIVEVGIRPGEKLYEELLVSTELV
DNQVMDKIFVGKVNVMPLESINQKIGEFRTLSGDELKQAIIAFANQTTHIEZ

ID101 1338bp

ATGATTGAACTTTATGATAGTTACAGTCAAGAAAGTCGAGATTTACATGAAAGTCTAGTCGCTACTGGTCTTTCTC
AACTTGGAGTGGTCATCGATGCAGATGGTTTCTGCCTGATGGTCTGCTTTCTCCTTTTACCTATTATCTAGGTTAC
GAGGATGGAAAACCTCTCTATTTTAATCAAGTTCCCGTTTCAGATTTTGGGAAATTTTAGGAGATAATCAGTCTG
CTTGTATTGAAGATGTGACGCAGGAGAGGGCTGTCATTCATTATGCTGATGGAATGCAGGCTCGCTTGGTTAAACA
GGTAGACTGGAAAGACCTAGAAGGTCGAGTACGTCAGGTTGACCACTACAATCGCTTCGGAGCTTGTTTTGCTAC
AACGACTTATAGCGCAGATAGCGAGCCGATTATGACAGTTTACCAAGATGTCAATGGTCAACAAGTTTTACTGGA
AAACCATGTGACGGGTGATATCTTATTGACTTTGCCAGGTCAGTCCATGCGTTACTTTGCAAATAAAGTTGAATTT
ATCACCTTCTTTTTGCAAGATTTGGAAATAGATACCAGTCAGCTTATCTTTAATACTCTAGCGACTCCTTTCTTGGT
TTCCTTCCATCATCCAGATAAATCTGGCTCGGATGTCTTGGTATGGCAGGAACCTCTCTATGATGCCATTCCAGGT

AATATGCAGTTGATTTTGGAAAGTGATAATGTGCGTACTAAGAAGATCATCATTCCAAATAAGGCGACTTATGAGC
GCGCTTTAGAGTTAACTGACGAGAAATACCATGATCAGTTTGTGCACTTGGGTTATCATTACCAGTTCAAACGTGA
TAATTTCCTAAGACGAGATGCCTTAATCTTGACCAATTCAGATCAGATTGAGCAAGTAGAAGCAATCGCAGGAGC
CTTGCCTGATGTCACTTTCCGTATTGCAGCGGTGACAGAGATGTCTTCTAAGCTCTTAGACATGCTTTGCTATCCTA
ATGTGGCCCTTTACCAGAACGCTAGTCCACAGAAGATTCAGGAGCTGTATCAACTGTCGGATATTTACTTGGATAT
AAACCACAGTAATGAGTTGCTACAGGCAGTGCGTCAGGCCTTTGAGCACAATCTCTTGATTCTTGGCTTTAATCAG
ACGGTGCACAATAGACTTTATATCGCTCCAGACCATCTATTTGAAAGTAGTGAAGTTGCTGCTTTGGTTGAGACCA
TTAAATTGGCCCTTTCAGATGTTGATCAAATGCGTCAGGCACTTGGCAAACAAGGCCAACATGCAAATTATGTTGA
CTTGGTGAGATATCAGGAAACCATGCAAACTGTTTTAGGAGGCTAA

MIELYDSYSQESRDLHESLVATGLSQLGVVIDADGFLPDGLLSPFTYYLGYEDGKPLYFNQVPVSDFWEILGDNQSACIE
DVTQERAVIHYADGMQARLVKQVDWKDLEGRVRQVDHYNRFGACFATTTYSADSEPIMTVYQDVNGQQVLLENHVT
GDILLTLPGQSMRYFANKVEFITFFLQDLEIDTSQLIFNTLATPFLVSFHHPDKSGSDVLVWQEPLYDAIPGNMQLILESD
NVRTKKIIIPNKATYERALELTDEKYHDQFVHLGYHYQFKRDNFLRRDALILTNSDQIEQVEAIAGALPDVTFRIAAVTE
MSSKLLDMLCYPNVALYQNASPQKIQELYQLSDIYLDINHSNELLQAVRQAFEHNLLILGFNQTVHNRLYIAPDHLFESS
EVAALVETIKLALSDVDQMRQALGKQGQHANYVDLVRYQETMQTVLGGZ

ID102 1512bp

ATGACAATTTACAATATAAATTTAGGAATTGGTTGGGCTAGTAGCGGTGTTGAATACGCTCAAGCCTATCGTGCTG
GTGTTTTTCGGAAATTAAATCTGTCCTCTAAGTTTATCTTTACAGATATGATTTTAGCCGATAATATTCAGCACTTA
ACAGCCAATATTGGTTTTGATGATAATCAGGTTATCTGGCTTTATAATCATTTCACAGATATCAAAATTGCACCTA
CTAGCGTGACAGTGGATGATGTCTTGGCTTACTTTGGTGGTGAAGAAAGTCACAGAGAAAAAAATGGCAAGGTTT
TACGTGTATTCTTTTTTGACCAAGATAAGTTTGTAACCTGTTATTTGGTTGATGAGAACAAGGACTTGGTTCAACAT
GCCGAGTATGTTTTTAAGGGAAACCTGATTCGGAAGGATTACTTTTCTTATACGCGTTATTGTAGCGAGTATTTTGC
TCCCAAGGACAATGTTGCAGTCTTATACCAACGAACTTTTTATAATGAAGACGGGACTCCAGTCTATGATATCTTG
ATGAATCAAGGGAAGGAAGAAGTTTATCATTTCAAGGATAAGATTTTCTATGGAAAGCAAGCTTTTGTGCGTGCCT
TTATGAAATCTTTGAATTTGAATAAGTCTGATTTGGTCATTCTCGATAGGGAGACAGGTATTGGACAGGTTGTGTT
TGAGGAAGCACAGACAGCACATCTAGCGGTAGTTGTTCATGCGGAGCATTATAGTGAAAATGCTACAAATGAGGA
CTATATCCTTTGGAATAACTATTATGACTATCAGTTTACCAATGCAGATAAGGTTGACTTCTTTATCGTGTCTACTG
ATAGACAAAATGAAGTTCTACAAGAGCAATTTGCCAAATATACTCAGCATCAGCCAAAGATTGTTACCATTCCTGT
AGGCAGTATTGATTCCTTGACAGATTCAAGTCAAGGGCGCAAACCATTTTCATTGATTACGGCTTCACGTCTTGCC
AAAGAAAAGCACATTGATTGGCTTGTGAAAGCTGTGATTGAAGCTCATAAGGAGTTACCGGAACTAACCTTTGAT
ATCTATGGTAGTGGTGGAGAAGATTCTCTGCTTAGAGAAATTATTGCAAATCATCAGGCAGAGGACTATATCCAAC
TCAAGGGGCATGCGGAACTTTCGCAGATTTATAGCCAGTATGAGGTCTACTTAACGGCTTCTACCAGCGAAGGATT
TGGTCTGACCTTGATGGAAGCTATTGGTTCAGGTCTACCTCTAATTGGTTTTGATGTGCCTTATGGTAATCAGACCT
TTATAGAGGATGGGCAAAATGGTTATTTGATTCCAAGTTCATCTGACCATGTAGAAGACCAAATCAAGCAAGCTTA
TGCCGCTAAGATTTGTCAATTGTATCAAGAAAATCGTTTGGAAGCTATGCGTGCCTATTCTTACCAAATTGCAGAA
GGCTTCTTGACCAAAGAAATTTTAGAAAAGTGGAAGAAAACAGTAGAGGAGGTGCTCCATGATTGA

MTIYNINLGIGWASSGVEYAQAYRAGVFRKLNLSSKFIFTDMILADNIQHLTANIGFDDNQVIWLYNHFTDIKIAPTSVT
VDDVLAYFGGEESHREKNGKVLRVFFFDQDKFVTCYLVDENKDLVQHAEYVFKGNLIRKDYFSYTRYCSEYFAPKDN
VAVLYQRTFYNEDGTPVYDILMNQGKEEVYHFKDKIFYGKQAFVRAFMKSLNLNKSDLVILDRETGIGQVVFEEAQTA
HLAVVVHAEHYSENATNEDYILWNNYYDYQFTNADKVDFFIVSTDRQNEVLQEQFAKYTQHQPKIVTIPVGSIDSLTDS
SQGRKPFSLITASRLAKEKHIDWLVKAVIEAHKELPELTFDIYGSGGEDSLLREIIANHQAEDYIQLKGHAELSQIYSQYE
VYLTASTSEGFGLTLMEAIGSGLPLIGFDVPYGNQTFIEDGQNGYLIPSSSDHVEDQIKQAYAAKICQLYQENRLEAMRA
YSYQIAEGFLTKEILEKWKKTVEEVLHDZ

ID103 2292bp

ATGTCCTCTCTTTCGGATCAAGAATTAGTAGCTAAAACAGTAGAGTTTCGTCAGCGTCTTTCCGAGGGAGAAAGTC
TAGACGATATTTTGGTTGAAGCTTTTGCTGTGGTGCGTGAAGCAGATAAGCGGATTTTAGGGATGTTTCCTTATGA
TGTTCAAGTCATGGGAGCTATTGTCATGCACTATGGAAATGTTGCTGAGATGAATACGGGGGAAGGTAAGACCTT
GACAGCTACCATGCCTGTCTATTTGAACGCTTTTTCAGGAGAAGGAGTGATGGTTGTGACTCCTAATGAGTATTTA
TCAAAGCGTGATGCCGAGGAAATGGGTCAAGTTTATCGTTTTCTAGGATTGACCATTGGTGTACCATTTACGGAAG
ATCCAAAGAAGGAGATGAAAGCTGAAGAAAAGAAGCTTATCTATGCTTCGGATATCATCTACACAACCAATAGTA
ATTTAGGTTTTGATTATCTAAATGATAACCTAGCCTCGAATGAAGAAGGTAAGTTTTTACGACCGTTTAACTATGT
GATTATTGATGAAATTGATGATATCTTGCTTGATAGTGCACAAACTCCTCTGATTATTGCGGGTTCTCCTCGTGTTC
AGTCTAATTACTATGCGATCATTGATACACTTGTAACAACCTTGGTCGAAGGAGAGGATTATATCTTTAAAGAGGA
GAAAGAGGAGGTTTGGCTCACTACTAAGGGGGCCAAGTCTGCTGAGAATTTCCTAGGGATTGATAATTTATACAA

GGAAGAGCATGCGTCTTTTGCTCGTCATTTGGTTTATGCGATTCGAGCTCATAAGCTCTTTACTAAAGATAAGGAC
TATATCATTCGTGGAAATGAGATGGTACTGGTTGATAAGGGAACAGGGCGTCTAATGGAAATGACTAAACTTCAA
GGAGGTCTCCATCAGGCTATTGAAGCCAAGGAACATGTCAAATTATCTCCTGAGACGCGGGCTATGGCCTCGATC
ACCTATCAGAGTCTTTTTAAGATGTTTAATAAGATATCTGGTATGACAGGGACAGGTAAGGTCGCGGAAAAAGAG
TTTATTGAAACTTACAATATGTCTGTAGTACGCATTCCAACCAATCGTCCGAGACAACGGATTGACTATCCAGATA
ATCTATATATCACTTTACCTGAAAAAGTGTATGCATCCTTGGAGTACATCAAGCAATACCATGCTAAGGGAAATCC
TTTACTCGTTTTTGTAGGCTCAGTTGAAATGTCTCAACTCTATTCGTCTCTCTTGTTTCGTGAAGGGATTGCCCATA
ATGTCCTAAATGCTAATAATGCGGCGCGTGAGGCTCAGATTATCTCCGAGTCAGGTCAGATGGGGGCTGTGACAG
TGGCTACCTCTATGGCAGGACGTGGTACGGATATCAAGCTTGGTAAAGGAGTCGCAGAGCTTGGGGGCTTGATTG
TTATTGGGACTGAGCGGATGGAAAGTCAGCGGATCGACCTACAAATTCGTGGCCGTTCTGGTCGTCAGGGAGATC
CTGGTATGAGTAAATTTTTTGTATCCTTAGAGGATGATGTTATCAAGAAATTTGGTCCATCTTGGGTGCATAAAAA
GTACAAAGACTATCAGGTTCAAGATATGACTCAACCGGAAGTATTGAAAGGTCGTAAATACCGGAAACTAGTCGA
AAAGGCTCAGCATGCCAGTGATAGTGCTGGACGTTCAGCACGTCGTCAGACTCTGGAGTATGCTGAAAGTATGAA
TATACAACGGGATATAGTCTATAAAGAGAGAAATCGTCTAATAGATGGTTCTCGTGACTTAGAGGATGTTGTTGTG
GATATCATTGAGAGATATACAGAAGAGGTAGCGGCTGATCACTATGCTAGTCGTGAATTATTGTTTCACTTTATTG
TGACCAATATTAGTTTTCATGTTAAAGAGGTTCCAGATTATATAGATGTAACTGACAAAACTGCAGTTCGTAGCTT
TATGAAGCAGGTGATTGATAAAGAACTTTCTGAAAAGAAAGAATTACTTAATCAACATGACTTATATGAACAGTTT
TTACGACTTTCACTGCTTAAAGCCATTGATGACAACTGGGTAGAGCAGGTAGACTATCTACAACAGCTATCCATGG
CTATCGGTGGTCAATCTGCTAGTCAGAAAAATCCAATCGTAGAGTACTATCAAGAAGCCTACGCGGGCTTTGAAG
CTATGAAAGAACAGATTCATGCGGATATGGTGCGTAATCTCCTGATGGGGCTGGTTGAGGTCACTCCAAAAGGTG
AAATCGTGACTCATTTTCCATAA

MSSLSDQELVAKTVEFRQRLSEGESLDDILVEAFAVVREADKRILGMFPYDVQVMGAIVMHYGNVAEMNTGEGKTLT
ATMPVYLNAFSGEGVMVVTPNEYLSKRDAEEMGQVYRFLGLTIGVPFTEDPKKEMKAEEKKLIYASDIIYTTNSNLGFD
YLNDNLASNEEGKFLRPFNYVIIDEIDDILLDSAQTPLIIAGSPRVQSNYYAIIDTLVTTLVEGEDYIFKEEKEEVWLTTKG
AKSAENFLGIDNLYKEEHASFARHLVYAIRAHKLFTKDKDYIIRGNEMVLVDKGTGRLMEMTKLQGGLHQAIEAKEHV
KLSPETRAMASITYQSLFKMFNKISGMTGTGKVAEKEFIETYNMSVVRIPTNRPRQRIDYPDNLYITLPEKVYASLEYIKQ
YHAKGNPLLVFVGSVEMSQLYSSLLFREGIAHNVLNANNAAREAQIISESGQMGAVTVATSMAGRGTDIKLGKGVAEL
GGLIVIGTERMESQRIDLQIRGRSGRQGDPGMSKFFVSLEDDVIKKFGPSWVHKKYKDYQVQDMTQPEVLKGRKYRKL
VEKAQHASDSAGRSARRQTLEYAESMNIQRDIVYKERNRLIDGSRDLEDVVVDIIERYTEEVAADHYASRELLFHFIVTN
ISFHVKEVPDYIDVTDKTAVRSFMKQVIDKELSEKKELLNQHDLYEQFLRLSLLKAIDDNWVEQVDYLQQLSMAIGGQS
ASQKNPIVEYYQEAYAGFEAMKEQIHADMVRNLLMGLVEVTPKGEIVTHFPZ

ID104 879bp

ATGAAACAAGAATGGTTTGAAAGTAATGATTTTGTAAAAACAACAAGCAAGAACAAGCCTGAAGAGCAAGCTCAA
GAGGTTGCAGACAAGGCTGAAGAAAGGATACCCGATCTCGATACACCAATTGAAAAAAATACTCAGTTAGAGGAG
GAAGTCTCTCAAGCTGAAGTCGAATTGGAAAGCCAGCAAGAAGAGAAAATTGAAGCTCCTGAAGACAGTGAAGC
GAGAACAGAAATAGAAGAAAAGAAGGCATCTAATTCTACTGAAGAAGAGCCAGACCTTTCTAAAGAAACAGAAA
AAGTCACTATAGCTGAAGAGAGCCAAGAAGCTCTTCCTCAGCAAAAAGCAACCACGAAAGAGCCACTTCTTATCA
GTAAATCTTTAGAAAGTCCTTATATCCCCGACCAAGCTCCAAAATCTAGGGATAAATGGAAAGAGCAAGTGCTTG
ATTTTGGTCTTGGCTAGTGGAAGCGATGAAATCTCCTACAAGTAAGTTGGAAACAAGTATCACACACAGTTACAC
AGCCTTTCTCTTGCTCATTCTGTTTTCTGCATCTTCCTTTTTCTTTAGTATCTATCACATCAAACATGCTTACTATGG
ACATATAGCAAGCATTAACAGTCGCTTCCCTGAGCAGCTAGCTCCTTTAACTCTTTTTCTATCATCTCTATCCTAG
TAGCGACAACACTCTTCTTCTTTTCATTCCTCTTGGGTAGTTTCGTTGTGAGACGATTTATCCACCAGGAAAAGGA
CTGGACGCTAGACAAGGTTCTCCAACAATATAGTCAACTCTTGGCAATTCCAATCTCCTCACTGCTATTGCTAGTT
TCTTTGCTTTCTTTGATAGCCTACGATTTACAGCCCTCTTGTGTGTGA

MKQEWFESNDFVKTTSKNKPEEQAQEVADKAEERIPDLDTPIEKNTQLEEEVSQAEVELESQQEEKIEAPEDSEARTEIE
EKKASNSTEEEPDLSKETEKVTIAEESQEALPQQKATTKEPLLISKSLESPYIPDQAPKSRDKWKEQVLDFWSWLVEAIKS
PTSKLETSITHSYTAFLLLILFSASSFFFSIYHIKHAYYGHIASINSRFPEQLAPLTLFSIISILVATTLFFFSFLLGSFVVRRFIH
QEKDWTLDKVLQQYSQLLAIPISSLLLLVSLLSLIAYDLQPSCVZ

ID106 327bp

ATGTAGTTTCCAACATGCTGTGCCTTGATTGAATTTCTCATCTTGGCTGTACTGGAGCAGGGTGATTCTTATGGTTA
TGAGATTAGCCAAACCATTAAGCTGATCGCTAATATCAAAGAATCCACACTCTATCCCATTCTCAAAAAATTGGAA
GGCAATAGCTTTCTGACAACCTATTCTAGAGAGTTCCAAGGTCGCATGCGCAAATACTACTCCTTGACAAACGGTG
GTATAGAGCAGCTCTTGACCCTAAAAGATGAATGGGCACTCTATACAGACACCATCAATGGCATCATAGAAGGGA
GTATCCGCCATGACAAGAACTGA

MYFPTSSALIEFLILAVLEQGDSYGYEISQTIKLIANIKESTLYPILKKLEGNSFLTTYSREFQGRMRKYYSLTNGGIEQLLT
LKDEWALYTDTINGIIEGSIRHDKNZ

ID108 954bp

ATGGATTTTGAAAAAATTGAACAAGCTTATATCTATTTACTAGAGAATGTCCAAGTCATCCAAAGTGATTTGGCGA
CCAACTTTTATGACGCCTTGGTGGAGCAAAATAGCATCTATCTGGATGGTGAAACTGAGCTAAACCAGGTCAAAG
ACAACAATCAGGCCCTTAAGCGTTTAGCACTACGCAAAGAAGAATGGCTCAAGACCTACCAGTTTCTCTTGATGA
AGGCTGGGCAAACAGAACCCTTGCAGGCCAATCACCAGTTTACACCGGATGCTATTGCTTTGCTTTTGGTGTTTAT
TGTGGAAGAGTTGTTTAAAGAGGAGGAAATTACTATCCTCGAAATGGGTTCTGGGATGGGAATTCTAGGCGCTATT
TTCTTGACCTCGCTTACTAAAAAGGTGGATTACTTGGGAATGGAAGTGGATGATTTGCTGATTGATCTGGCAGCTA
GCATGGCAGATGTAATTGGTTTGCAGGCTGGCTTTGTCCAAGGAGATGCCGTTCGCCCACAAATGCTCAAAGAAA
GCGATGTGGTCATCAGTGACTTGCCTGTCGGCTATTATCCTGATGATGCCGTTGCGTCGCGCCATCAAGTTGCTTC
TAGCCAAGAACATACTTACGCCCATCACTTGCTCATGGAACAAGGGCTTAAGTACCTCAAGTCAGACGGATACGC
TATTTTTCTAGCTCCGAGTGATTTGTTGACCAGTCCTCAAAGTGATTTGTTAAAAGAATGGCTGAAAGAAGAGGCG
AGTCTGGTTGCTATGATTAGTCTGCCTGAAAATCTCTTTGCTAATGCCAAACAATCTAAGACTATTTTTATCTTACA
GAAGAAAAATGAAATAGCAGTAGAGCCTTTTGTTTATCCACTTGCTAGCTTGCAAGATGCAAGTGTTTTAATGAAA
TTTAAAGAAAATTTTCAAAAATGGACTCAAGGTACTGAAATATAA

MDFEKIEQAYIYLLENVQVIQSDLATNFYDALVEQNSIYLDGETELNQVKDNNQALKRLALRKEEWLKTYQFLLMKAG
QTEPLQANHQFTPDAIALLLVFIVEELFKEEEITILEMGSGMGILGAIFLTSLTKKVDYLGMEVDDLLIDLAASMADVIGL
QAGFVQGDAVRPQMLKESDVVISDLPVGYYPDDAVASRHQVASSQEHTYAHHLLMEQGLKYLKSDGYAIFLAPSDLLT
SPQSDLLKEWLKEEASLVAMISLPENLFANAKQSKTIFILQKKNEIAVEPFVYPLASLQDASVLMKFKENFQKWTQGTEI
Z

ID110 1902bp

ATGATTATTTTACAAGCTAATAAAATTGAACGTTCTTTTGCAGGAGAGGTTCTTTTCGATAATATCAACCTGCAGG
TTGATGAACGAGATCGGATTGCTCTTGTTGGGAAAAATGGTGCAGGTAAGTCTACTCTTTTGAAGATTTTAGTTGG
AGAAGAGGAGCCAACTAGCGGAGAAATCAATAAGAAAAAAGATATTTCTCTGTCTTACCTAGCCCAAGATAGCCG
TTTTGAGTCTGAAAATACCATCTACGATGAAATGCTTCATGTCTTTAATGATTTGCGTCGGACGGAGAGACAACTG
CGTCAGATGGAGCTGGAGATGGGTGAAAAGTCTGGTGAGGATTTGGATAAACTGATGTCAGATTATGACCGCTTA
TCTGAGAATTTTCGCCAAGCAGGTGGCTTTACCTATGAAGCTGATATTCGAGCGATTTTGAATGGATTCAAGTTTG
ACGAGTCTATGTGGCAGATGAAAATTGCTGAGCTTTCTGGTGGTCAAAATACTCGTTTGGCACTTGCCAAAATGCT
CCTTGAAAAGCCCAATCTCTTGGTCTTGGACGAGCCAACTAACCACTTGGATATTGAAACCATCGCCTGGCTAGAG
AATTACTTGGTAAACTATAGCGGTGCCCTCATTATCGTCAGCCACGACCGTTATTTCTTGGACAAGGTTGCGACAA
TTACGCTAGATTTGACCAAGCATTCCTTGGATCGCTATGTGGGGAATTACTCTCGTTTTGTCGAATTGAAGGAGCA
AAAGCTAGTTACTGAGGCAAAAAACTATGAAAAGCAACAGAAGGAAATCGCTGCTCTGGAAGACTTTGTCAATCG
CAATCTAGTTCGTGCTTCAACGACTAAACGTGCTCAATCTCGCCGTAAACAACTAGAAAAAATGGAGCGTTTGGA
CAAGCCTGAAGCTGGCAAGAAAGCAGCCAACATGACCTTCCAGTCTGAAAAAACGTCGGGCAATGTTGTTTTGAC
TGTTGAAAATGCAGCTGTTGGCTATGACGGGGAAGTCTTGTCACAACCTATCAACCTAGATCTTCGTAAGATGAAT
GCTGTCGCTATCGTTGGTCCAAATGGTATCGGCAAGTCAACCTTTATCAAGTCTATTGTGGACCAGATTCCTTTTAT
CAAGGGAGAAAAGCGCTTTGGCGCTAATGTTGAGGTTGGTTACTATGACCAAACCCAAAGCAAGCTGACACCAAG
TAATACGGTGCTGGATGAACTCTGGAATGATTTCAAACTGACACCAGAAGTTGAAATCCGCAACCGTCTTGGAGC
CTTCCTTTTCTCAGGAGATGATGTTAAAAAATCAGTCGGCATGCTATCTGGTGGCGAAAAAGCTCGTTTGCTTTTA
GCTAAATTGTCTATGGAAAACAATAACTTTTTGATTCTGGATGAGCCGACCAACCACTTGGATATTGATAGTAAGG
AAGTGCTAGAAAATGCCTTGATTGACTTTGATGGAACCTTGCTGTTTGTCAGTCATGATCGTTACTTTATCAATCGT
GTGGCAACTCATGTTTTGGAATTGTCTGAGAATGGTTCAACTCTCTACCTTGGAGATTACGACTACTATGTTGAGA
AGAAAGCAACAGCAGAAATGAGTCAGACTGAGGAAGCTTCAACTAGCAATCAAGCAAAGGAAGCAAGTCCAGTC
AATGACTATCAGGCCCAGAAAGAAAGTCAAAAAGAAGTTCGCAAACTCATGCGACAAATCGAAAGTCTAGAAGCT
GAAATTGAAGAGCTAGAAAGTCAAAGCCAAGCCATTTCTGAACAAATGTTGGAAACAAACGATGCCGACAAACTC
ATGGAATTACAGGCTGAGCTGGACAAAATCAGCCATCGTCAGGAAGAAGCTATGCTTGAGTGGGAAGAATTATCA
GAGCAGGTGTAA

MIILQANKIERSFAGEVLFDNINLQVDERDRIALVGKNGAGKSTLLKILVGEEEPTSGEINKKKDISLSYLAQDSRFESENT
IYDEMLHVFNDLRRTERQLRQMELEMGEKSGEDLDKLMSDYDRLSENFRQAGGFTYEADIRAILNGFKFDESMWQMK
IAELSGGQNTRLALAKMLLEKPNLLVLDEPTNHLDIETIAWLENYLVNYSGALIIVSHDRYFLDKVATITLDLTKHSLDR
YVGNYSRFVELKEQKLVTEAKNYEKQQKEIAALEDFVNRNLVRASTTKRAQSRRKQLEKMERLDKPEAGKKAANMTF
QSEKTSGNVVLTVENAAVGYDGEVLSQPINLDLRKMNAVAIVGPNGIGKSTFIKSIVDQIPFIKGEKRFGANVEVGYYDQ
TQSKLTPSNTVLDELWNDFKLTPEVEIRNRLGAFLFSGDDVKKSVGMLSGGEKARLLLAKLSMENNNFLILDEPTNHLD
IDSKEVLENALIDFDGTLLFVSHDRYFINRVATHVLELSENGSTLYLGDYDYYVEKKATAEMSQTEEASTSNQAKEASP
VNDYQAQKESQKEVRKLMRQIESLEAEIEELESQSQAISEQMLETNDADKLMELQAELDKISHRQEEAMLEWEELSEQV
Z

ID111 1179bp

ATGAATCGCTATGCAGTGCAGTTGATTAGCCGTGGGGCTATCAATAAAATGGGAAATATGCTCTATGATTATGGAA
ATAGTGTCTGGTTGGCTTCTATGGGGACTATAGGACAGACAGTTTTAGGAATGTATCAGATTTCTGAGCTCGTCAC
ATCTATTCTCGTCAATCCCTTTGGCGGAGTTATTTCAGACCGTTTTTCTCGTCGTAAGATTTTAATGACGGCAGATC
TTGTTTGTGGGATTCTTTGTCTGGCTATTTCTTTCATAAGGAATGATAGCTGGATGATTGGCGCTTTGATTGTTGCT
AACATTGTGCAGGCTATTGCTTTTGCCTTTTCTCGCACAGCCAATAAAGCTATCATAACTGAAGTGGTGGAGAAAG
ATGAGATTGTGATCTATAATTCTCGCTTAGAGCTGGTTTTGCAGGTTGTAGGTGTTAGCTCTCCTGTTCTTTCCTTC
CTTGTTTTACAGTTTGCAAGTCTCCATATGACGCTACTGCTAGACTCGCTGACTTTTTCATTGCTTTTGTTCTAGT
GGCTTTCCTTCCAAAAGAGGAAGCAAAAGTTCAAGAGAAAAAGGCTTTTACTGGGAGAGATATTTTTGTAGATAT
CAAGGATGGGTTACACTATATCTGGCATCAGCAAGAAATTTTCTTCCTTTTGCTGGTAGCTTCCAGCGTTAATTTCT
TTTTTGCAGCTTTTGAATTTCTACTTCCCTTTTCGAATCAGCTTTACGGGTCAGAAGGAGCCTATGCAAGTATTTTA
ACTATGGGGGCTATTGGTTCCATCATTGGGGCTCTTCTAGCTAGTAAAATTAAAGCTAATATTTATAATCTTTTGAT
TTTACTGGCTTTGACAGGTGTCGGAGTTTTTATGATGGGATTACCACTTCCAACTTTTCTTTCCTTTTCTGGAAATT
TAGTTTGTGAATTGTTTATGACGATTTTTAATATTCACTTTTTTACTCAAGTACAAACCAAGGTTGAGAGCGAATTT
CTTGGAAGAGTACTGAGTACAATTTTTACCTTAGCTATTCTATTTATGCCTATTGCAAAAGGATTTATGACAGTCTT
GCCAAGTGTCCATCTTTATTCTTTCTTGATTATTGGACTTGGAGTTGTAGCCTTATATTTCTTAGCTCTCGGATATG
TTCGAACTCATTTTGAAAAATTGATATAA

MNRYAVQLISRGAINKMGNMLYDYGNSVWLASMGTIGQTVLGMYQISELVTSILVNPFGGVISDRFSRRKILMTADLV
CGILCLAISFIRNDSWMIGALIVANIVQAIAFAFSRTANKAIITEVVEKDEIVIYNSRLELVLQVVGVSSPVLSFLVLQFASL
HMTLLLDSLTFFIAFVLVAFLPKEEAKVQEKKAFTGRDIFVDIKDGLHYIWHQQEIFFLLLVASSVNFFFAAFEFLLPFSN
QLYGSEGAYASILTMGAIGSIIGALLASKIKANIYNLLILLALTGVGVFMMGLPLPTFLSFSGNLVCELFMTIFNIHFFTQV
QTKVESEFLGRVLSTIFTLAILFMPIAKGFMTVLPSVHLYSFLIIGLGVVALYFLALGYVRTHFEKLIZ

ID113 2466bp

ATGCAAAATCAATTAAATGAATTAAAACGAAAAATGCTGGAATTTTTCCAGCAAAAACAAAAAAATAAAAAATCA
GCTAGACCTGGCAAGAAAGGTTCAAGTACCAAAAAATCTAAAACCTTAGATAAGTCAGCCATTTTCCCAGCTATTT
TACTGAGTATAAAAGCCTTATTTAACTTACTCTTTGTACTCGGTTTTCTAGGAGGAATGTTGGGAGCTGGGATTGCT
TTGGGATACGGAGTGGCCTTATTTGACAAGGTTCGGGTGCCTCAGACAGAAGAATTGGTGAATCAGGTCAAGGAC
ATCTCTTCTATTTCAGAGATTACCTATTCGGACGGGACGGTGATTGCTTCCATAGAGAGTGATTTGTTGCGCACTTC
TATCTCATCTGAGCAAATTTCGGAAAATCTGAAGAAGGCTATCATTGCGACAGAAGATGAACACTTTAAAGAACA
TAAGGGTGTAGTACCCAAGGCGGTGATTCGTGCGACCTTGGGGAAATTTGTAGGTTTGGGTTCCTCTAGTGGGGGT
TCAACCTTGACCCAGCAACTAATTAAACAGCAGGTGGTTGGGGATGCGCCGACCTTGGCTCGTAAGGCGGCAGAG
ATTGTGGATGCTCTTGCCTTGGAACGCGCCATGAATAAAGATGAGATTTTAACGACCTATCTCAATGTGGCTCCCT
TTGGCCGAAATAATAAGGGACAGAATATTGCAGGGGCTCGGCAAGCAGCTGAGGGAATTTTCGGTGTAGATGCCA
GTCAGTTGACTGTTCCTCAAGCAGCATTTTTAGCAGGACTTCCACAGAGTCCCATTACTTACTCTCCTTATGAAAA
TACTGGGGAGTTGAAGAGTGATGAAGACCTAGAAATTGGCTTAAGACGGGCTAAGGCAGTTCTTTACAGTATGTA
TCGTACAGGTGCATTAAGCAAAGACGAGTATTCTCAGTACAAGGATTATGACCTTAAACAGGACTTTTTACCATCG
GGCACGGTTACAGGAATTTCACGAGACTATTTATACTTTACAACTTTGGCAGAAGCTCAAGAACGTATGTATGACT
ATCTAGCTCAGAGAGACAATGTCTCCGCTAAGGAGTTGAAAAATGAGGCAACTCAGAAGTTTTATCGAGATTTGG
CAGCCAAGGAAATTGAAAATGGTGGTTATAAGATTACTACTACCATAGATCAGAAAATTCATTCTGCCATGCAAA
GTGCGGTTGCTGATTATGGCTATCTTTTAGACGATGGAACAGGTCGTGTAGAAGTAGGGAATGTCTTGATGGATAA
CCAAACAGGTGCTATTCTAGGCTTTGTAGGTGGTCGTAATTATCAAGAAAATCAAAATAATCATGCCTTTGATACC
AAACGTTCGCCAGCTTCTACTACCAAGCCCTTGCTGGCCTACGGTATTGCTATTGACCAGGGCTTGATGGGAAGTG
AAACGATTCTATCTAACTATCCAACAAAGTTTGCTAATGGCAATCCGATTATGTATGCTAATAGCAAGGGAACAGG
AATGATGACCTTGGGAGAAGCTCTGAACTATTCATGGAATATCCCTGCTTACTGGACCTATCGTATGCTCCGTGAA
AAGGGTGTTGATGTCAAGGGTTATATGGAAAAGATGGGTTACGAGATTCCTGAGTACGGTATTGAGAGCTTGCCA
ATGGGTGGTGGTATTGAAGTCACAGTTGCCCAGCATACCAATGGCTATCAGACCTTAGCTAATAATGGAGTTTATC
ATCAGAAGCATGTGATTTCAAAGATTGAAGCAGCAGATGGTAGAGTGGTGTATGAGTATCAGGATAAACCGGTTC
AAGTCTATTCAAAAGCTACTGCGACGATTATGCAGGGATTGCTACGAGAAGTTCTATCCTCTCGTGTGACAACAAC
CTTCAAGTCTAACCTGACTTCTTTAAATCCTACTCTGGCTAATGCAGATTGGATTGGGAAGACTGGTACAACCAAC
CAAGACGAAAATATGTGGCTCATGCTTTCGACACCTAGATTAACCCTAGGTGGCTGGATTGGGCATGATGATAATC
ATTCATTGTCACGTAGAGCAGGTTATTCTAATAACTCTAATTACATGGCTCATCTGGTAAATGCGATTCAGCAAGC
TTCCCCAAGCATTTGGGGGAACGAGCGCTTTGCTTTAGATCCTAGTGTAGTGAAATCGGAAGTCTTGAAATCAACA
GGTCAAAAACCAGAGAAGGTTTCTGTTGAAGGAAAAGAAGTAGAGGTCACAGGTTCGACTGTTACCAGCTATTGG
GCTAATAAGTCAGGAGCGCCAGCGACAAGTTATCGCTTTGCTATTGGCGGAAGTGATGCGGATTATCAGAATGCTT
GGTCTAGTATTGTGGGGAGTCTACCAACTCCATCCAGCTCCAGCAGTTCAAGTAGTAGTTCTAGCGATAGCAGTAA
CTCAAGTACTACACGACCTTCTTCTTCAAGGGCGAGACGATAA

MQNQLNELKRKMLEFFQQKQKNKKSARPGKKGSSTKKSKTLDKSAIFPAILLSIKALFNLLFVLGFLGGMLGAGIALGY
GVALFDKVRVPQTEELVNQVKDISSISEITYSDGTVIASIESDLLRTSISSEQISENLKKAIIATEDEHFKEHKGVVPKAVIRA
TLGKFVGLGSSSGGSTLTQQLIKQQVVGDAPTLARKAAEIVDALALERAMNKDEILTTYLNVAPFGRNNKGQNIAGAR
QAAEGIFGVDASQLTVPQAAFLAGLPQSPITYSPYENTGELKSDEDLEIGLRRAKAVLYSMYRTGALSKDEYSQYKDYD
LKQDFLPSGTVTGISRDYLYFTTLAEAQERMYDYLAQRDNVSAKELKNEATQKFYRDLAAKEIENGGYKITTTIDQKIH

SAMQSAVADYGYLLDDGTGRVEVGNVLMDNQTGAILGFVGGRNYQENQNNHAFDTKRSPASTTKPLLAYGIAIDQGL
MGSETILSNYPTNFANGNPIMYANSKGTGMMTLGEALNYSWNIPAYWTYRMLREKGVDVKGYMEKMGYEIPEYGIES
LPMGGGIEVTVAQHTNGYQTLANNGVYHQKHVISKIEAADGRVVYEYQDKPVQVYSKATATIMQGLLREVLSSRVTTT
FKSNLTSLNPTLANADWIGKTGTTNQDENMWLMLSTPRLTLGGWIGHDDNHSLSRRAGYSNNSNYMAHLVNAIQQAS
PSIWGNERFALDPSVVKSEVLKSTGQKPEKVSVEGKEVEVTGSTVTSYWANKSGAPATSYRFAIGGSDADYQNAWSSIV
GSLPTPSSSSSSSSSSDSSNSSTTRPSSSRARRZ

ID114 1974bp

ATGAAAAAATTTTATGTAAGTCCAATTTTTCCTATTCTAGTAGGATTGATTGCGTTTGGAGTCTTATCCACTTTCAT
TATTTTTGTTAATAATAATCTGTTGACGGTTTTAATTTTGTTTCTTTTTGTAGGAGGCTATGTTTTTTTATTTAAGAA
ACTGAGAGTGCATTATACAAGGAGTGATGTAGAACAGATACAGTATGTAAACCACCAAGCGGAAGAAAGTTTGAC
AGCTCTATTGGAACAGATGCCTGTAGGTGTTATGAAATTGAATTTATCTTCTGGAGAGGTTGAGTGGTTTAATCCC
TATGCTGAATTGATTTTGACCAAGGAAGATGGTGATTTTGATTTAGAAGCTGTTCAAACGATTATCAAGGCTTCAG
TAGGAAATCCGTCTACTTATGCCAAGCTTGGTGAGAAGCGTTATGCTGTTCATATGGATGCTTCTTCCGGTGTTTTG
TATTTTGTAGATGTATCCAGGGAACAAGCCATAACAGATGAATTGGTAACAAGTAGACCAGTGATTGGGATTGTCT
CTGTGGATAATTATGATGATTTGGAGGATGAAACTTCTGAGTCAGATATTAGTCAAATCAATAGTTTTGTAGCTAA
TTTTATATCAGAGTTTTCAGAAAAACACATGATGTTTTCTCGTCGGGTAAGTATGGATCGATTTTATCTATTTACTG
ACTACACGGTGCTTGAGGGCTTGATGAATGATAAATTTTCTGTTATTGATGCTTTCAGAGAAGAGTCGAAACAGAG
ACAGTTGCCCTTGACCTTAAGTATGGGATTTTCTTATGGCGATGGAAATCATGATGAGATAGGGAAAGTTGCTTTG
CTCAATTTGAACTTGGCTGAAGTACGTGGTGGCGACCAGGTGGTTGTTAAGGAAAACGACGAAACGAAAAATCCA
GTTTATTTTGGTGGTGGGTCTGCTGCTTCAATCAAGCGTACACGGACTCGTACGCGCGCTATGATGACAGCTATTT
CAGATAAGATTCGGAGTGTAGATCAGGTTTTTGTAGTCGGTCACAAAAATTTAGACATGGATGCTTTGGGCTCTGC
TGTAGGTATGCAGTTGTTCGCCAGCAATGTGATTGAAAATAGCTATGCTCTTTATGATGAAGAACAAATGTCTCCA
GATATTGAACGAGCTGTTTCATTCATAGAAAAAGAAGGAGTTACGAAGTTGTTGTCTGTTAAGGATGCAATGGGG
ATGGTGACCAATCGTTCTTTGTTGATTCTTGTAGACCATTCAAAGACAGCCTTAACATTATCAAAAGAATTTTATG
ATTTATTTACCCAAACCATTGTTATTGACCACCATAGAAGGGATCAGGATTTTCCAGATAATGCGGTTATTACTTA
TATCGAAAGTGGTGCAAGTAGTGCCAGTGAGTTGGTAACGGAATTGATTCAGTTCCAGAATTCTAAGAAAAATCG
TTTGAGTCGTATGCAAGCAAGTGTCTTGATGGCTGGTATGATGTTGGATACTAAAAATTTCACCTCGCGAGTAACT
AGTCGGACATTTGATGTTGCTAGCTATCTCAGAACGCGCGGAAGTGATAGTATTGCTATCCAGGAAATCGCTGCGA
CAGATTTTGAAGAATATCGTGAGGTCAATGAACTTATTTTACAGGGGCGTAAATTAGGTTCAGATGTACTAATAGC
AGAGGCTAAGGACATGAAATGCTATGATACAGTTGTTATTAGTAAGGCAGCAGATGCCATGTTAGCCATGTCAGG
TATTGAAGCGAGTTTTGTTCTTGCGAAGAATACACAAGGATTTATCTCTATCTCAGCTCGAAGTCGTAGTAAACTG
AATGTACAACGGATTATGGAAGAGTTAGGCGGTGGAGGCCACTTTAATTTGGCAGCAGCTCAAATTAAAGATGTA
ACCTTGTCAGAAGCAGGTGAAAAACTGACAGAAATTGTATTAAATGAAATGAAGGAAAAGGAGAAAGAAGAATG
A

MKKFYVSPIFPILVGLIAFGVLSTFIIFVNNNLLTVLILFLFVGGYVFLFKKLRVHYTRSDVEQIQYVNHQAEESLTALLEQ
MPVGVMKLNLSSGEVEWFNPYAELILTKEDGDFDLEAVQTIIKASVGNPSTYAKLGEKRYAVHMDASSGVLYFVDVSR
EQAITDELVTSRPVIGIVSVDNYDDLEDETSESDISQINSFVANFISEFSEKHMMFSRRVSMDRFYLFTDYTVLEGLMNDK
FSVIDAFREESKQRQLPLTLSMGFSYGDGNHDEIGKVALLNLNLAEVRGGDQVVVKENDETKNPVYFGGGSAASIKRTR
TRTRAMMTAISDKIRSVDQVFVVGHKNLDMDALGSAVGMQLFASNVIENSYALYDEEQMSPDIERAVSFIEKEGVTKL
LSVKDAMGMVTNRSLLILVDHSKTALTLSKEFYDLFTQTIVIDHHRRDQDFPDNAVITYIESGASSASELVTELIQFQNSK
KNRLSRMQASVLMAGMMLDTKNFTSRVTSRTFDVASYLRTRGSDSIAIQEIAATDFEEYREVNELILQGRKLGSDVLIAE
AKDMKCYDTVVISKAADAMLAMSGIEASFVLAKNTQGFISISARSRSKLNVQRIMEELGGGGHFNLAAAQIKDVTLSEA
GEKLTEIVLNEMKEKEKEEZ

ID115 663bp

ATGAAGTGCTTGTTATGTGGCAGACTATGAAGACTGTTTTAACTTTTAGTAGTCTCTTACTTCTGAGGAATGATG
ACTCTTGTCTTTGTTCAGACTGTGATTCTACTTTTGAAAGAATTGGGGAAGAGAACTGTCCAAATTGTATGAAAAC
AGAGTTGTCAACAAAGTGTCAAGATTGTCAACTTTGGTGTAAAGAGGGAGTTGAAGTCAGTCATAGAGCGATTTTT
ACTTACAATCAAGCTATGAAGGATTTTTTCAGTCGGTATAAGTTTGATGGAGACTTCCTGTTAAGAAAAGTTTTCG
CTTCATTTTTAAGTGAGGAGTTGAAAAAGTACAAAGAGTATCAATTTGTTGTAATTCCCCTAAGTCCTGATAGATA
TGCTAATAGAGGATTTAATCAGGTTGAGGGCTTGGTAGAGGCAGCAGGCTTTGAGTATCTGGATTTATTAGAGAAA
AGAGAAGAGAGAGCCAGTTCTTCTAAAAATCGTTCAGAGCGCTTGCGGGACAGAACTTCCTTTCTTTATTAAAAGTG
GAGTCACTATTCCTAAAAAAATCCTACTTATAGATGATATCTATACTACAGGAGCAACTATAAATCGTGTTAAGAA
ACTGTTGGAAGAAGCTGGTGCTAAGGATGTAAAAACATTTTCCCTTGTAAGATGA

MKCLLCGQTMKTVLTFSSLLLLRNDDSCLCSDCDSTFERIGEENCPNCMKTELSTKCQDCQLWCKEGVEVSHRAIFTY
NQAMKDFFSRYKFDGDFLLRKVFASFLSEELKKYKEYQFVVIPLSPDRYANRGFNQVEGLVEAAGFEYLDLLEKREER
ASSSKNRSERLGTELPFFIKSGVTIPKKILLIDDIYTTGATINRVKKLLEEAGAKDVKTFSLVRZ

ID116 1299bp

ATGAAAGTAAATTTAGATTATCTCGGTCGTTTATTTACTGAGAATGAATTAACAGAAGAAGAACGTCAGTTGGCGG
AGAAACTTCCAGCAATGAGAAAGGAGAAGGGGAAACTTTTCTGTCAACGCTGTAATAGTACTATTCTAGAAGAAT
GGTATTTGCCCATCGGTGCTTACTATTGTCGAGAGTGCTTGCTGATGAAGCGAGTCAGAAGTGATCAAACTTTATA
CTATTTTCCGCAGGAGGATTTTCCAAAGCAAGATGTTCTCAAATGGCGCGGCCAATTAACTCCTTTTCAAGAGAAG
GTGTCAGAGGGATTGCTTCAAGTAGTAGACAAGCAAAAGCCAACCTTAGTTCATGCGGTAACAGGAGCTGGAAAG
ACAGAAATGATTTATCAAGTAGTGGCTAAAGTGATCAATGCGGGTGGTGCAGTGTGTTTGGCTAGTCCTCGCATAG
ATGTTTGTTTGGAGCTGTACAAGCGCCTGCAACAGGATTTTTCTTGCGGGATAGCTTTGCTACATGGAGAATCGGA
ACCTTATTTTCGAACACCACTAGTTGTTGCAACAACCCATCAGTTATTGAAGTTTTATCAAGCTTTTGATTTGCTGA
TAGTGGATGAAGTAGATGCTTTTCCTTATGTTGATAATCCCATGCTTTACCACGCTGTCAAGAATAGTGTAAAGGA
GAATGGATTGAGAATCTTTTTAACAGCGACTTCGACCAATGAGTTAGATAAAAAGGTCCGTTTAGGAGAACTAAA
AAGACTGAATTTACCGAGACGGTTTCATGGAAATCCGTTGATTATTCCAAAACCAATTTGGTTATCGGATTTTAAT
CGCTACTTAGACAAGAATCGTTTGTCACCAAAGTTAAAGTCCTATATTGAGAAGCAGAGAAAGACAGCTTATCCG
TTACTCATTTTTGCTTCAGAAATTAAGAAAGGGGAGCAGTTAGCAGAAATCTTACAGGAGCAATTTCCAAATGAG
AAAATTGGCTTTGTATCTTCTGTAACAGAGGATCGATTAGAGCAAGTACAAGCTTTTCGAGATGGAGAACTGACA
ATACTTATCAGTACGACAATCTTGGAGCGCGGAGTTACCTTCCCTTGTGTGGATGTTTTCGTAGTAGAGGCCAATC
ATCGTTTGTTTACCAAGTCTAGTTTGATTCAGATTGGTGGACGAGTTGGACGAAGCATGGATAGACCGACAGGAG
ATTTGCTTTTCTTCCATGATGGGTTAAATGCTTCAATCAAGAAGGCGATTAAGGAAATTCAGATGATGAATAAGGA
GGCTGGTCTATGA

MKVNLDYLGRLFTENELTEEERQLAEKLPAMRKEKGKLFCQRCNSTILEEWYLPIGAYYCRECLLMKRVRSDQTLYYF
PQEDFPKQDVLKWRGQLTPFQEKVSEGLLQVVDKQKPTLVHAVTGAGKTEMIYQVVAKVINAGGAVCLASPRIDVCLE
LYKRLQQDFSCGIALLHGESEPYFRTPLVVATTHQLLKFYQAFDLLIVDEVDAFPYVDNPMLYHAVKNSVKENGLRIFL
TATSTNELDKKVRLGELKRLNLPRRFHGNPLIIPKPIWLSDFNRYLDKNRLSPKLKSYIEKQRKTAYPLLIFASEIKKGEQL
AEILQEQFPNEKIGFVSSVTEDRLEQVQAFRDGELTILISTTILERGVTFPCVDVFVVEANHRLFTKSSLIQIGGRVGRSMD
RPTGDLLFFHDGLNASIKKAIKEIQMMNKEAGLZ

ID117 870bp

ATGCAAATTCAAAAAAGTTTTAAGGGGCAGTCTCCCTATGGCAAGCTGTATCTAGTGGCAACGCCGATTGGCAAT
CTAGATGATATGACTTTTCGTGCTATCCAGACCTTGAAAGAAGTGGACTGGATTGCTGCTGAGGATACGCGCAATA
CAGGGCTTTTGCTCAAGCATTTTGACATTTCCACCAAGCAGATCAGTTTTCATGAGCACAATGCCAAGGAAAAAAT
TCCTGATTTGATTGGTTTCTTGAAAGCAGGGCAAAGTATTGCTCAGGTCTCTGATGCCGGTTTGCCTAGCATTTCA
GACCCTGGTCATGATTTAGTTAAGGCAGCTATTGAGGAAGAAATTGCAGTTGTGACAGTTCCAGGTGCCTCTGCAG
GAATTTCTGCCTTGATTGCCAGTGGTTTAGCGCCACAGCCACATATCTTTTACGGTTTTTTACCGAGAAAATCAGG
TCAGCAGAAGCAATTTTTTGGCTTGAAAAAAGATTATCCTGAAACACAGATTTTTTATGAATCACCTCATCGTGTA
GCAGACACGTTGGAAAATATGTTAGAAGTCTACGGTGACCGCTCCGTTGTCTTGGTCAGGGAATTGACCAAAATCT
ATGAAGAATACCAACGAGGTACTATCTCTGAGTTATTAGAAAGCATTGCTGAAACGCCACTCAAGGGCGAATGTC
TTCTCATTGTTGAGGGTGCCAGTCAGGGTGTGGAGGAAAAGGACGAGGAAGACTTGTTCGTAGAAATTCAAACCC
GCATCCAGCAAGGTGTGAAGAAAAACCAAGCTATCAAGGAAGTCGCTAAGATTTACCAGTGGAATAAAAGTCAGC
TCTACGCTGCCTACCACGACTGGGAAGAAAAACAATAA

MQIQKSFKGQSPYGKLYLVATPIGNLDDMTFRAIQTLKEVDWIAAEDTRNTGLLLKHFDISTKQISFHEHNAKEKIPDLI
GFLKAGQSIAQVSDAGLPSISDPGHDLVKAAIEEEIAVVTVPGASAGISALIASGLAPQPHIFYGFLPRKSGQQKQFFGLKK
DYPETQIFYESPHRVADTLENMLEVYGDRSVVLVRELTKIYEEYQRGTISELLESIAETPLKGECLLIVEGASQGVEEKDE
EDLFVEIQTRIQQGVKKNQAIKEVAKIYQWNKSQLYAAYHDWEEKQZ

ID118 345bp

ATGATAAAGAAAGGAAAGGGCTGTTTTATGGACAAAAAAGAATTATTTGACGCGCTGGATGATTTTTCCCAACAA
TTATTGGTAACCTTAGCCGATGTGGAAGCCATCAAGAAAAATCTCAAGAGCCTGGTAGAGGAAAATACAGCTCTT
CGCTTGGAAAATAGTAAGTTGCGAGAACGCTTGGGTGAGGTGGAAGCAGATGCTCCTGTCAAGGCCAAGCATGTT
CGCGAAAGTGTCCGTCGTATTTACCGTGATGGATTTCACGTATGTAATGATTTTTATGGACAACGTCGAGAGCAGG
ACGAAGAATGTATGTTTTGTGACGAGTTGTTATACAGGGAGTAA

MIKKGKGCFMDKKELFDALDDFSQQLLVTLADVEAIKKNLKSLVEENTALRLENSKLRERLGEVEADAPVKAKHVRES
VRRTYRDGFHVCNDFYGQRREQDEECMFCDELLYREZ

ID119 639bp

ATGTCAAAAGGATTTTTAGTCTCTCTTGAGGGACCAGAGGGAGCAGGCAAGACCAGTGTTTTAGAGGCTCTGCTAC
CAATTTTAGAGGAAAAAGGAGTAGAGGTGTTGACGACCCGTGAACCTGGCGGAGTCTTGATTGGGGAGAAGATTC
GGGAAGTGATTTTGGATCCAAGTCATACTCAGATGGATGCTAAAACAGAGCTACTTCTCTATATTGCCAGTCGCAG
ACAGCATTTGGTGGAAAAAGTTCTTCCAGCCCTTGAAGCTGGCAAGTTGGTCATCATGGATCGTTTTATCGATAGT

TCTGTTGCCTATCAGGGATTTGGTCGTGGCTTAGATATTGAAGCCATTGACTGGCTCAATCAGTTTGCGACAGATG
GCCTCAAACCCGATTTGACACTCTATTTTGACATCGAGGTGGAAGAAGGGCTGGCTCGTATTGCTGCTAATAGTGA
CCGCGAGGTTAATCGTTTGGATTTGGAAGGGTTGGACTTGCATAAAAAAGTTCGTCAAGGCTACCTTTCTCTTCTG
GATAAAGAGGGAAATCGCATTGTCAAGATTGATGCTAGTCTCCCTTTGGAGCAAGTTGTGGAAACTACCAAGGCT
GTCTTGTTTGACGGAATGGGCTTGGCCAAATGA

MSKGFLVSLEGPEGAGKTSVLEALLPILEEKGVEVLTTREPGGVLIGEKIREVILDPSHTQMDAKTELLLYIASRRQHLVE
KVLPALEAGKLVIMDRFIDSSVAYQGFGRGLDIEAIDWLNQFATDGLKPDLTLYFDIEVEEGLARIAANSDREVNRLDLE
GLDLHKKVRQGYLSLLDKEGNRIVKIDASLPLEQVVETTKAVLFDGMGLAKZ

ID120 408bp

ATGGTAGAACAAAGAAAATCAATTACCATGAAAGATGTTGCTTTAGAAGCAGGAGTTAGTGTTGGAACTGTTTCA
CGTGTAATTAATAAAGAAAAAGGCATTAAAGAAGTAACTTTGAAAAAAGTGGAACAAGCGATTAAAACTTTGAAT
TACATTCCAGATTACTACGCTAGAGGAATGAAAAAAAATCGAACAGAAACGATTGCAATCATTGTACCAAGTATC
TGGCATCCCTTCTTTTCAGAATTTGCTATGCATGTGGAAAATGAAGTCTATAAGAGAAATAACAAATTACTCTTAT
GTTCTATCAATGGTACAAATAGAGAGCAAGACTATCTGGAGATGTTGCGTCATAATAAAGTTGATGGAGTGGTTGC
CATTACCTATAGGCCAATTGAACATTACTTGACGTCAGGAATTCCCTTTGTTAGTATTGACCGCACATACTCAGAG
ATTGCCATTCCTTGTGTTTCA

MVEQRKSITMKDVALEAGVSVGTVSRVINKEKGIKEVTLKKVEQAIKTLNYIPDYYARGMKKNRTETIAIIVPSIWHPFFS
EFAMHVENEVYKRNNKLLLCSINGTNREQDYLEMLRHNKVDGVVAITYRPIEHYLTSGIPFVSIDRTYSEIAIPCVS

ID121 285bp

ATGAATATATTTAGAACAAAGAATGTTAGTTTAGATAAAACAGAGATGCATAGGCATTTGAAGTTATGGGATTTGA
TTTTGCTGGGTATCGGAGCCATGGTAGGGACAGGCGTCTTTACAATCACAGGTACTGCAGCTGCAACACTTGCTGG
CCCAGCCCTAGTGATTTCAATCGTTATTTCTGCCTTGTGTGTGGGATTATCAGCCCTCTTTTTTGCAGAATTTGCCT
CGCGAGTACCCGCTACAGGAGGTGCCTATAGTTACCTCTATGCTATCTTAGGAGAATTCCCTGCCTGGTTGGCTGG
TTGGTTAACCATGATGGAGTTCATGACAGCCATATCAGGCGTAGCTTCGGGTTGGGCAGCTTATTTTTAA

MNIFRTKNVSLDKTEMHRHLKLWDLILLGIGAMVGTGVFTITGTAAATLAGPALVISIVISALCVGLSALFFAEFASRVP
ATGGAYSYLYAILGEFPAWLAGWLTMMEFMTAISGVASGWAAYF

ID124 1311bp

ATGAAATCAAGAGTAAAGGAAACGAGTATGGATAAAATTGTGGTTCAAGGTGGCGATAATCGTCTGGTAGGAAGC
GTGACGATCGAGGGAGCAAAAAATGCAGTCTTACCCTTGTTGGCAGCGACTATTCTAGCAAGTGAAGGAAAGACC
GTCTTGCAGAATGTTCCGATTTTGTCGGATGTCTTTATTATGAATCAGGTAGTTGGTGGTTTGAATGCCAAGGTTGA
CTTTGATGAGGAAGCTCATCTTGTCAAGGTGGATGCTACTGGCGACATCACTGAGGAAGCCCCTTACAAGTATGTC
AGCAAGATGCGCGCCTCCATCGTTGTATTAGGGCCAATCCTTGCCCGTGTGGGTCATGCCAAGGTATCCATGCCAG
GTGGTTGTACGATTGGTAGCCGTCCTATTGATCTTCATTTGAAAGGTCTGGAAGCTATGGGGGTTAAGATTAGTCA
GACAGCTGGTTACATCGAAGCCAAGGCAGAACGCTTGCATGGTGCTCATATCTATATGGACTTTCCAAGTGTTGGT
GCAACGCAGAACTTGATGATGGCAGCGACTCTGGCTGATGGGGTGACAGTGATTGAGAATGCTGCGCGTGAGCCT
GAGATTGTTGACTTAGCCATTCTCCTTAATGAAATGGGAGCCAAGGTCAAAGGTGCTGGTACAGAGACTATAACC
ATTACTGGTGTTGAGAAACTTCATGGTACGACTCACAATGTAGTCCAAGACCGTATCGAAGCAGGAACCTTTATGG
TAGCTGCTGCCATGACTGGTGGTGATGTCTTGATTCGAGACGCTGTCTGGGAGCACAACCGTCCCTTGATTGCCAA
GTTACTTGAAATGGGTGTTGAAGTAATTGAAGAAGACGAAGGAATTCGTGTTCGTTCTCAACTAGAAAATCTAAA
AGCTGTTCATGTGAAAACCTTGCCCCACCCAGGATTTCCAACAGATATGCAGGCTCAATTTACAGCCTTGATGACA
GTTGCAAAAGGCGAATCAACCATGGTGGAGACAGTTTTCGAAAATCGTTTCCAAACCTAGAAGAGATGCGCCGCA
TGGGCTTGCATTCTGAGATTATCCGTGATACAGCTCGTATTGTTGGTGGACAGCCTTTGCAGGGAGCAGAAGTTCT
TTCAACTGACCTTCGTGCCAGTGCGGCCTTGATTTTGACAGGTTTGGTAGCACAGGGAGAAACTGTGGTCGGTAAA
TTGGTTCACTTGGATAGAGGTTACTACCGTTTCCATGAGAAGTTGGCGCAGCTAGGTGCTAAGATTCAGCGGATTG
AGGCAAGTGATGAAGATGAATAA

MKSRVKETSMDKIVVQGGDNRLVGSVTIEGAKNAVLPLLAATILASEGKTVLQNVPILSDVFIMNQVVGGLNAKVDFD
EEAHLVKVDATGDITEEAPYKYVSKMRASIVVLGPILARVGHAKVSMPGGCTIGSRPIDLHLKGLEAMGVKISQTAGYIE
AKAERLHGAHIYMDFPSVGATQNLMMAATLADGVTVIENAAREPEIVDLAILLNEMGAKVKGAGTETITITGVEKLHG
TTHNVVQDRIEAGTFMVAAAMTGGDVLIRDAVWEHNRPLIAKLLEMGVEVIEEDEGIRVRSQLENLKAVHVKTLPHPG
FPTDMQAQFTALMTVAKGESTMVETVFENRFQHLEEMRRMGLHSEIIRDTARIVGGQPLQGAEVLSTDLRASAALILTG
LVAQGETVVGKLVHLDRGYYGFHEKLAQLGAKIQRIEASDEDEZ

ID125 1101bp

ATGTTATTAGCGTCAACAGTAGCCTTGTCATTTGCCCCAGTATTGGCAACTCAAGCAGAAGAAGTTCTTTGGACTG
CACGTAGTGTTGAGCAAATCCAAAACGATTTGACTAAAACGGACAACAAAACAAGTTATACCGTACAGTATGGTG
ATACTTTGAGCACCATTGCAGAAGCCTTGGGTGTAGATGTCACAGTGCTTGCGAATCTGAACAAAATCACTAATAT
GGACTTGATTTTCCCAGAAACTGTTTTGACAACGACTGTCAATGAAGCAGAAGAAGTAACAGAAGTTGAAATCCA
AACACCTCAAGCAGACTCTAGTGAAGAAGTGACAACTGCGACAGCAGATTTGACCACTAATCAAGTGACCGTTGA
TGATCAAACTGTTCAGGTTGCAGACCTTTCTCAACCAATTGCAGAAGTTACAAAGACAGTGATTGCTTCTGAAGAA
GTGGCACCATCTACGGGCACTTCTGTCCCAGAGGAGCAAACGACCGAAACAACTCGCCCAGTTGCAGAAGAAGCT
CCTCAGGAAACGACTCCAGCTGAGAAGCAGGAAACACAAACAAGCCCTCAAGCTGCATCAGCAGTGGAAGCAAC
TACAACAAGTTCAGAAGCAAAAGAAGTAGCATCATCAAATGGAGCTACAGCAGCAGTTTCTACTTATCAACCAGA
AGAAACGAAAGTAATTTCAACAACTTACGAGGCTCCAGCTGCGCCCGATTATGCTGGACTTGCAGTAGCAAAATC
TGAAAATGCAGGTCTTCAACCACAAACAGCTGCCTTTAAWGAAGAAATTGCTAACTTGTTTGGCATTACATCCTTT
AGTGGTTATCGTCCAGGAGACAGTGGAGATCACGGAAAAGGTTTGGCTATCGACTTTATGGTACCAGAACGTTCA
GAATTAGGGGATAAGATTGCGGAATATGCTATTCAAAATATGGCCAGCCGTGGCATTAGTTACATCATCTGGAAA
CAACGTTTCTATGCTCCATTCGATAGCAAATATGGGCCAGCTAACACTTGGAACCCAATGCCAGACCGTGGTAGTG
TGACAGAAAATCACTATGATCACGTTCACGTTTCAATGAATGGATAA

MLLASTVALSFAPVLATQAEEVLWTARSVEQIQNDLTKTDNKTSYTVQYGDTLSTIAEALGVDVTVLANLNKITNMDL
IFPETVLTTTVNEAEEVTEVEIQTPQADSSEEVTTATADLTTNQVTVDDQTVQVADLSQPIAEVTKTVIASEEVAPSTGTS
VPEEQTTETTRPVAEEAPQETTPAEKQETQTSPQAASAVEATTSSEAKEVASSNGATAAVSTYQPEETKVISTTYEAPA
APDYAGLAVAKSENAGLQPQTAAFKKKLLTCLALHPLVVIVQETVEITEKVWLSTLWYQNVQNZGIRLRNMLFKIWPA
VALVTSSGNNVSMLHSIANMGQLTLGTQCQTVVVZQKITMITFTFQZMD

ID126 1281bp

TTGTTTAAGAAAAAATAAAGACATTCTTAATATTGCATTGCCAGCTATGGGTGAAAACTTTTTGCAGATGCTAATGG
GAATGGTGGACAGTTATTTGGTTGCTCATTTAGGATTGATAGCTATTTCAGGGGTTTCAGTAGCTGGTAATATTAT
CACCATTTATCAGGCGATTTTCATCGCTCTGGGAGCTGCTATTTCCAGTGTTATTTCAAAAAGCATAGGGCAGAAA
GACCAGTCGAAGTTGGCCTATCATGTGACTGAGGCGTTGAAGATTACCTTACTATTAAGTTTCCTTTTAGGATTTTT
GTCCATCTTCGCTGGGAAAGAGATGATAGGACTTTTGGGGACGGAGAGGGATGTAGCTGAGAGTGGTGGACTGTA
TCTATCTTTGGTAGGCGGATCGATTGTTCTCTTAGGTTTAATGACTAGTCTAGGAGCCTTGATTCGTGCAACGCATA
ATCCACGTCTGCCTCTCTATGTTAGTTTTTATCCAATGCCTTGAATATTCTTTTTTCAAGTCTAGCTATTTTTGTTC
TGGATATGGGGATAGCTGGTGTTGCTTGGGGGACAATTGTGTCTCGTTTGGTTGGTCTTGTGATTTTGTGGTCACA
ATTAAAACTGCCTTATGGGAAGCCAACTTTTGGTTTAGATAAGGAACTGTTGACCTTGGCTTTACCAGCAGCTGGA
GAGCGACTTATGATGAGGGCTGGAGATGTAGTGATCATTGCCTTGGTCGTTTCTTTTGGGACGGAGGCAGTTGCTG
GGAATGCAATCGGAGAAGTCTTGACCCAGTTTAACTATATGCCTGCCTTTGGCGTCGCTACGGCAACGGTCATGCT
GTTGGCCCGAGCAGTTGGAGAGGATGATTGGAAAAGAGTTGCTAGTTTGAGTAAACAAACCTTTTGGCTTTCTCTG
TTCCTCATGTTGCCCCTGTCCTTTAGTATATATGTCTTGGGTGTACCATTAACTCATCTCTATACGACTGATTCTCT
AGCGGTGGAGGCTAGTGTTCTAGTGACACTGTTTTCACTACTTGGGACCCCTATGACGACAGGAACAGTCATCTAT
ACGGCAGTCTGGCAGGGATTAGGAAATGCACGCCTCCCTTTTTATGCGACAAGTATAGGAATGTGGTGTATCCGC
ATTGGGACAGGATATCTGATGGGGATTGTGCTTGGTTGGGGCTTGCCTGGTATTTGGGCAGGGTCTCTCTTGGATA
ATGGTTTTCGCTGGTTATTTCTACGCTATCGTTACCAGCGCTATATGAGCTTGAAAGGATAG

LFKKNKDILNIALPAMGENFLQMLMGMVDSYLVAHLGLIAISGVSVAGNIITIYQAIFIALGAAISSVISKSIGQKDQSKLA
YHVTEALKITLLLSFLLGFLSIFAGKEMIGLLGTERDVAESGGLYLSLVGGSIVLLGLMTSLGALIRATHNPRLPLYVSFL
SNALNILFSSLAIFVLDMGIAGVAWGTIVSRLVGLVILWSQLKLPYGKPTFGLDKELLTLALPAAGERLMMRAGDVVIIA
LVVSFGTEAVAGNAIGEVLTQFNYMPAFGVATATVMLLARAVGEDDWKRVASLSKQTFWLSLFLMLPLSFSIYVLGVP
LTHLYTTDSLAVEASVLVTLFSLLGTPMTTGTVIYTAVWQGLGNARLPFYATSIGMWCIRIGTGYLMGIVLGWGLPGIW
AGSLLDNGFRWLFLRYRYQRYMSLKGZ

ID127 894bp

GTGGGAAGAATTATCAGAGCAGGTGTAAAGATGGAACATCTTGGAAAAGTATTTTCGTGAATTTCGAACAAGTGGA
AATTATTCTTTAAAGGAAGCAGCAGGCGAATCCTGCTCTACCTCTCAGTTATCTCGCTTTGAGCTTGGGGAGTCTG
ACCTGGCAGTCTCCCGTTTCTTTGAGATTTTGGATAACATTCATGTAACAATCGAAAATTTCATGGATAAGGCAAG
GAATTTTCATAATCATGAACATGTGTCTATGATGGCACAGATTATCCCACTTTACTATTCAAACGATATTGCAGGT
TTTGAAAAGCTTCAAAGAGAACAACTTGAAAAGTCTAAGAGTTCGACGACTCCCCTTTATTTTGAGCTGAACTGGA
TTTTGCTACAAGGTCTGATTTGTCAAAGAGATGCGAGTTATGATATGAAGCAGGATGATTTGGGTAAGGTAGCAGA
TTATCTCTTCAAAACAGAAGAATGGACCATGTATGAGTTGATTCTTTTCGGTAACCTCTATAGTTTCTACGATGTA
GACTATGTCACTCGGATTGGTAGAGAAGTTATGGAGAGGCAGGAATTTTACCAAGAGATTAGTCGCCATAAGAGA
TTAGTGTTGATTTTGGCCCTCAATTGTTACCAGCATTGTTTAGAGCATTCTTCTTTTATAATGCCAACTATTTTGA

GGCTTATACAGAGAAGATTATTGACAAAGGTATTAAGCTTTATGAGCGTAATGTTTTCCATTATTTAAAAGGTTTT
GCCTTATATCAAAAAGGACAGTGTAAAGAAGGCTGTAAGCAGATGCAAGAGGCCATGCATATTTTTGATGTGTTA
GGTCTTCCAGAGCAAGTAGCCTATTATCAGGAACACTACGAAAAATTTGTCAAAAGTTAA

VGRIIRAGVKMEHLGKVFREFRTSGNYSLKEAAGESCSTSQLSRFELGESDLAVSRFFEILDNIHVTIENFMDKARNFHN
HEHVSMMAQIIPLYYSNDIAGFQKLQREQLEKSKSSTTPLYFELNWILLQGLICQRDASYDMKQDDLGKVADYLFKTEE
WTMYELILFGNLYSFYDVDYVTRIGREVMEREEFYQEISRHKRLVLILALNCYQHCLEHSSFYNANYFEAYTEKIIDKGI
KLYERNVFHYLKGFALYQKGQCKEGCKQMQEAMHIFDVLGLPEQVAYYQEHYEKFVKSZ

TABLE 3

ID1 1068bp

ATGTCTAACATTCAAAACATGTCCCTGGAGGACATCATGGGAGAGCGCTTTGGTCGCTACTCCAAGTACATTATTC
AAGACCGGGCTTTGCCAGATATTCGTGATGGGTTGAAGCCGGTTCAGCGCCGTATTCTTTATTCTATGAATAAGGA
TAGCAATACTTTTGACAAGAGCTACCGTAAGTCGGCCAAGTCAGTCGGGAACATCATGGGGAATTTCCACCCACA
CGGGGATTCTTCTATCTATGATGCCATGGTTCGTATGTCACAGAACTGGAAAAATCGTGAGATTCTAGTTGAAATG
CACGGTAATAACGGTTCTATGGACGGAGATCCTCCTGCGGCTATGCGTTATACTGAGGCACGTTTGTCTGAAATTG
CAGGCTACCTTCTTCAGGATATCGAGAAAAAGACAGTTCCTTTTGCATGGAACTTTGACGATACGGAGAAAGAAC
CAACGGTCTTGCCAGCAGCCTTTCCAAACCTCTTGGTCAATGGTTCGACTGGGATTTCGGCTGGTTATGCCACAGA
CATTCCTCCCCATAATTTAGCTGAGGTCATAGATGCTGCAGTTTACATGATTGACCACCCAACTGCAAAGATTGAT
AAACTCATGGAATTCTTGCCTGGACCAGACTTCCCTACAGGGGCTATTATTCAGGGTCGTGATGAAATCAAGAAA
GCTTATGAGACTGGGAAAGGGCGCGTGGTTGTTCGTTCCAAGACTGAAATTGAAAAGCTAAAAGGTGGTAAGGAA
CAAATCGTTATTATTGAGATTCCTTATGAAATCAATAAGGCCAATCTAGTCAAGAAAATCGATGATGTTCGTGTTA
ATAACAAGGTAGCTGGGATTGCTGAGGTTCGTGATGAGTCTGACCGTGATGGTCTTCGTATCGCTATCGAACTTAA
GAAAGACGCTAATACTGAGCTTGTTCTCAACTACTTATTTAAGTACACCGACCTACAAATCAACTACAACTTTAAT
ATGGTGGCGATTGACAATTTCACACCTCGTCAGGTTGGATTGTTCCAATCCTGTCTAGCTATATCGCTCACCGTCG
AGAAGTGA

MSNIQNMSLEDIMGERFGRYSKYIIQDRALPDIRDGLKPVQRRILYSMNKDSNTFDKSYRKSAKSVGNIMGNFHPHGDSS
IYDAMVRMSQNWKNREILVEMHGNNGSMDGDPPAAMRYTEARLSEIAGYLLQDIEKKTVPFAWNFDDTEKEPTVLPA
AFPNLLVNGSTGISAGYATDIPPHNLAEVIDAAVYMIDHPTAKIDKLMEFLPGPDFPTGAIIQGRDEIKKAYETGKGRVV
VRSKTEIEKLKGGKEQIVIIEIPYEINKANLVKKIDDVRVNNKVAGIAEVRDESDRDGLRIAIELKKDANTELVLNYLFKY
TDLQINYNFNMVAIDNFTPRQVGLFQSCLAISLTVEKZ

ID12 684bp

ATGCCGACATTAGAAATAGCACAAAAAAAACTGGAGTTCATTAAGAAGGCAGAAGAATATTACAATGCCTTGTGT
ACAAATATACAGTTGAGCGGAGATAAACTAAAAGTAATTTCCGTTACTTCTGTTAACCCTGGGGAAGGAAAAACA
ACTACTTCCATAAATATAGCATGGTCGTTTGCGCGTGCAGGCTATAAAACTCTTTTGATCGATGGCGATACTCGAA
ATTCAGTTATGTTAGGAGTTTTTAAATCTCGTGAAAAAATTACAGGGCTAACAGAATTTTTATCTGGGACAGCTGA
TTTATCTCACGGTTTATGTGATACAAATATTGAAAATTTATTTGTAGTTCAATCGGGATCTGTATCACCAAACCCTA
CAGCCTTGTTACAAAGTAAAAATTTTAATGATATGATTGAAACATTGCGTAAATATTTTGATTATATCATTATTGAT
ACACCGCCTATTGGAATTGTTATTGATGCGGCAATTATCACTCAAAAGTGTGATGCGTCCATCTTGGTAACAGCAA
CAGGTGAGGCGAATAAACGTGATATCCAAAAAGCGAAACAACAATTAAAACAAACAGGGAAACTGTTCCTAGGA
GTTGTTTTAAATAAATTGGATATCTCGGTTAATAAGTATGGAGTTTACGGTTCCTATGGAAATTATGGTAAAAAAT
AA

MPTLEIAQKKLEFIKKAEEYYNALCTNIQLSGDKLKVISYTSVNPGEGKTTTSINIAWSFARAGYKTLLIDGDTRNSVML
GVFKSREKITGLTEFLSGTADLSHGLCDTNIENLFVVQSGSVSPNPTALLQSKNFNDMIETLRKYFDYIIIDTPPIGIVIDAA
IITQKCDASILVTATGEANKRDIQKAKQQLKQTGKLFLGVVLNKLDISVNKYGVYGSYGNYGKKZ

ID13 1182bp

ATGGAGGCAAATATGAAACATCTAAAAACATTTTACAAAAAATGGTTTCAATTATTAGTCGTTATCGTCATTAGCT
TTTTTAGTGGAGCCTTGGGTAGTTTTTCAATAACTCAACTAACTCAAAAAAGTAGTGTAAACAACTCTAACAACAA
TAGTACTATTACACAAACTGCCTATAAGAACGAAAATTCAACAACACAGGCTGTTAACAAAGTAAAAGATGCTGT
TGTTTCTGTTATTACTTATTCGGCAAACAGACAAAATAGCGTATTTGGCAATGATGATACTGACACAGATTCTCAG
CGAATCTCTAGTGAAGGATCTGGAGTTATTTATAAAAAGAATGATAAAGAAGCTTACATCGTCACCAACAATCAC
GTTATTAATGGCGCCAGCAAAGTAGATATTCGATTGTCAGATGGGACTAAAGTACCTGGAGAAATTGTCGGAGCT
GACACTTTCTCTGATATTGCTGTCGTCAAAATCTCTTCAGAAAAAGTGACAACAGTAGCTGAGTTTGGTGATTCTA
GTAAGTTAACTGTAGGAGAAACTGCTATTGCCATCGGTAGCCCGTTAGGTTCTGAATATGCAAATACTGTCACTCA
AGGTATCGTATCCAGTCTCAATAGAAATGTATCCTTAAAATCGGAAGATGGACAAGCTATTTCTACAAAAGCCATC
CAAACTGATACTGCTATTAACCCAGGTAACTCTGGCGGCCCACTGATCAATATTCAAGGGCAGGTTATCGGAATTA
CCTCAAGTAAAATTGCTACAAATGGAGGAACATCTGTAGAAGGTCTTGGTTTCGCAATTCCTGCAAATGATGCTAT
CAATATTATTGAACAGTTAGAAAAAAACGGAAAAGTGACGCGTCCAGCTTTGGGAATCCAGATGGTTAATTTATC
TAATGTGAGTACAAGCGACATCAGAAGACTCAATATTCCAAGTAATGTTACATCTGGTGTAATTGTTCGTTCGGTA
CAAAGTAATATGCCTGCCAATGGTCACCTTGAAAAATACGATGTAATTACAAAAGTAGATGACAAAGAGATTGCT
TCATCAACAGACTTACAAAGTGCTCTTTACAACCATTCTATCGGAGACACCATTAAGATAAGCTACTATCGTAAGG
GGAAAGAAGAAACTACCTCTATCAAACTTAACAAGAGTTCAGGTGATTTAGAATCTTAA

MEANMKHLKTFYKKWFQLLVVIVISFFSGALGSFSITQLTQKSSVNNSNNNSTITQTAYKNENSTTQAVNKVKDAVVSV
ITYSANRQNSVFGNDDTDTDSQRISSEGSGVIYKKNDKEAYIVTNNHVINGASKVDIRLSDGTKVPGEIVGADTFSDIAVV

KISSEKVTTVAEFGDSSKLTVGETAIAIGSPLGSEYANTVTQGIVSSLNRNVSLKSEDGQAISTKAIQTDTAINPGNSGGPLI
NIQGQVIGITSSKIATNGGTSVEGLGFAIPANDAINIIEQLEKNGKVTRPALGIQMVNLSNVSTSDIRRLNIPSNVTSGVIVR
SVQSNMPANGHLEKYDVITKVDDKEIASSTDLQSALYNHSIGDTIKITYYRNGKEETTSIKLNKSSGDLESZ

ID15 939bp

ATGGCAGAAATTTATCTAGCAGGTGGTTGTTTTTGGGGCCTAGAGGAATATTTTTCACGCATTTCTGGAGTGCTAG
AAACCAGTGTTGGCTACGCTAATGGTCAAGTCGAAACGACCAATTACCAGTTGCTCAAGGAAACAGACCATGCAG
AAACGGTCCAAGTGATTTACGATGAGAAGGAAGTGTCACTCAGAGAGATTTTACTTTATTATTTCCGAGTTATCGA
TCCTCTATCTATCAATCAACAAGGGAATGACCGTGGTCGCCAATATCGAACTGGGATTTATTATCAGGATGAAGCA
GATTTGCCAGCTATCTACACAGTGGTGCAGGAGCAGGAACGCATGCTGGGTCGAAAGATTGCAGTAGAAGTGGAG
CAATTACGCCACTACATTCTGGCTGAAGACTACCACCAAGACTATCTCAGGAAGAATCCTTCAGGTTACTGTCATA
TCGATGTGACCGATGCTGATAAGCCATTGATTGATGCAGCAAACTATGAAAAGCCTAGTCAAGAGGTGTTGAAGG
CCAGTCTATCTGAAGAGTCTTATCGTGTCACACAAGAAGCTGCTACAGAGGCTCCATTTACCAATGCCTATGACCA
AACCTTTGAAGAGGGGATTTATGTAGATATTACGACAGGTGAGCCACTCTTTTTTGCCAAGGATAAGTTTGCTTCA
GGTTGTGGTTGGCCAAGTTTTAGCCGTCCGATTTCCAAAGAGTTGATTCATTATTACAAGGATCTGAGCCATGGAA
TGGAGCGAATTGAAGTTCGTTCTCGTTCAGGCAGTGCTCACTTGGGTCATGTTTTCACAGATGGACCGCGGGAGTT
AGGCGGCCTCCGTTACTGTATCAATTCTGCTTCTTTACGCTTTGTGGCCAAGGATGAGATGGAAAAAGCAGGATAT
GGCTATCTATTGCCTTACTTAAACAAATAA

MAEIYLAGGCFWGLEEYFSRISGVLETSVGYANGQVETTNYQLLKETDHAETVQVIYDEKEVSLREILLYYFRVIDPLSI
NQQGNDRGRQYRTGIYYQDEADLPAIYTVVQEQERMLGRKIAVEVEQLRHYILAEDYHQDYLRKNPSGYCHIDVTDA
DKPLIDAANYEKPSQEVLKASLSEESYRVTQEAATEAPFTNAYDQTFEEGIYVDITTGEPLFFAKDKFASGCGWPSFSRPI
SKELIHYYKDLSHGMERIEVRSRSGSAHLGHVFTDGPRELGGLRYCINSASLRFVAKDEMEKAGYGYLLPYLNKZ

ID17 870bp

ATGAAGATTATTGTACCTGCAACCAGTGCCAATATCGGGCCAGGTTTTGACTCGGTCGGTGTAGCTGTAACCAAGT
ATCTTCAAATTGAGGTCTGCGAAGAACGAGATGAGTGGCTGATTGAACACCAGATTGGCAAATGGATTCCACATG
ACGAGCGTAATCTCTTGCTCAAAATCGCTTTGCAAATTGTACCAGACTTGCAACCAAGACGCTTGAAAATGACCAG
TGATGTCCCTTTGGCGCGCGGTTTGGGTTCTTCCAGCTCGGTTATCGTTGCTGGGATTGAACTAGCCAACCAACTG
GGTCAACTCAACTTATCAGACCATGAAAAATTGCAGTTAGCGACCAAGATTGAAGGGCATCCTGACAATGTGGCT
CCAGCCATTTATGGTAATCTCGTTATTGCAAGTTCTGTTGAAGGGCAAGTCTCTGCTATCGTAGCAGACTTTCCAG
AGTGTGATTTTCTAGCTTACATTCCAAACTATGAATTACGTACTCGCGACAGCCGTAGTGTCTTGCCTAAAAAATT
GTCTTATAAGGAAGCTGTTGCTGCAAGTTCTATCGCCAATGTAGCGGTTGCTGCCTTGTTGGCAGGAGACATGGTG
ACCGCTGGGCAAGCAATCGAGGGAGACCTCTTCCATGAGCGCTATCGTCAGGACTTGGTAAGAGAATTTGCGATG
ATTAAGCAAGTGACCAAAGAAAATGGGGCCTATGCAACCTACCTTTCTGGTGCTGGGCCGACAGTTATGGTTCTG
GCTTCTCATGACAAGATGCCAACAATTAAGGCAGAATTGGAAAAGCAACCTTTCAAAGGAAAACTGCATGACTTG
AGAGTTGATACCCAAGGTGTCCGTGTAGAAGCAAAATAA

MKIIVPATSANIGPGFDSVGVAVTKYLQIEVCEERDEWLIEHQIGKWIPHDERNLLLKIALQIVPDLQPRRLKMTSDVPLA
RGLGSSSSVIVAGIELANQLGQLNLSDHEKLQLATKIEGHPDNVAPAIYGNLVIASSVEGQVSAIVADFPECDFLAYIPNY
ELRTRDSRSVLPKKLSYKEAVAASSIANVAVAALLAGDMVTAGQAIEGDLFHERYRQDLVREFAMIKQVTKENGAYAT
YLSGAGPTVMVLASHDKMPTIKAELEKQPFKGKLHDLRVDTQGVRVEAKZ

ID20 564bp

ATGAAATATCACGATTACATCTGGGATTTAGGTGGAACTTTACTGGATAATTATGAAACTTCAACAGCTGCATTTG
TTGAAACATTGGCACTGTATGGTATCACACAAGACCATGACAGTGTCTATCAAGCTTTAAAGGTTTCTACTCCTTT
TGCGATTGAGACATTCGCTCCCAATTTAGAGAATTTTTTAGAAAAGTACAAGGAAAATGAAGCCAGAGAGCTTGA
ACACCCGATTTTATTTGAAGGAGTTTCTGACCTATTGGAAGACATTTCAAATCAAGGTGGCCGTCATTTTTTGGTCT
CTCATCGAAATGATCAGGTTTTGGAAATTTTAGAAAAAACCTCTATAGCAGCTTATTTTACAGAAGTGGTGACTTC
TAGCTCAGGCTTTAAGAGAAAGCCAAATCCCGAATCCATGCTTTATTTAAGAGAAAAGTATCAGATTAGCTCTGGT
CTTGTCATTGGTGATCGGCCGATTGATATCGAAGCAGGTCAAGCTGCAGGACTTGATACCCACTTGTTTACCAGTA
TCGTGAATTTAAGACAAGTATTAGACATATAA

MKYHDYIWDLGGTLLDNYETSTAAFVETLALYGITQDHDSVYQALKVSTPFAIETFAPNEENFLEKYKENEARELEHPI
LFEGVSDLLEDISNQGGRHFLVSHRNDQVLEILEKTSIAAYFTEVVTSSSGFKRKPNPESMLYLREKYQISSGLVIGDRPID
IEAGQAAGLDTHLFTSIVNLRQVLDIZ

ID21 1875bp

ATGACAGAAGAAATCAAAAATCTGCAGGCACAGGATTATGATGCCAGTCAAATTCAAGTTTTAGAGGGCTTAGAG
GCTGTTCGTATGCGTCCAGGGATGTACATTGGATCAACCTCAAAAGAAGGTCTTCACCATCTAGTCTGGGAAATTG

TTGATAACTCAATTGACGAGGCCTTGGCAGGATTTGCCAGCCATATTCAAGTTTTTATTGAGCCAGATGATTCGAT
TACTGTTGTGGATGATGGGCGTGGTATCCCAGTCGATATTCAGGAAAAAACAGGCCGTCCTGCTGTTGAGACCGTC
TTTACAGTCCTTCACGCTGGAGGAAAGTTCGGCGGTGGTGGATACAAGGTTTCAGGTGGTCTTCACGGGGTGGGGT
CGTCAGTAGTTAATGCCCTTTCCACTCAATTAGACGTTCATGTTCACAAAAATGGTAAGATTCATTACCAAGAATA
CCGTCGTGGTCATGTTGTCGCAGATCTTGAAATAGTTGGAGATACGGATAAAACAGGAACAACTGTTCACTTCACA
CCGGACCCAAAAATCTTCACTGAAACAACAATCTTTGATTTTGATAAATTAAATAAACGGATTCAAGAGTTGGCCT
TTCTAAATCGCGGTCTTCAAATTTCAATTACAGATAAGCGCCAAGGTTTGGAACAAACCAAGCATTATCATTATGA
AGGTGGGATTGCTAGTTACGTTGAATATATCAACGAGAACAAGGATGTAATCTTTGATACACCAATCTATACAGAC
GGTGAGATGGATGATATCACAGTTGAGGTAGCCATGCAGTACACAACTGGTTACCATGAAAATGTCATGAGTTTC
GCCAATAATATTCATACCCATGAAGGTGGAACACATGAACAAGGTTTCCGTACAGCCTTGACACGTGTTATCAAC
GATTATGCTCGTAAAAATAAGTTACTGAAAGACAATGAAGATAATTTAACAGGGGAAGATGTTCGCGAAGGCTTA
ACTGCAGTTATCTCAGTTAAACACCCAAATCCACAGTTTGAAGGACAAACCAAGACCAAATTGGGAAATAGCGAA
GTGGTCAAGATTACCAATCGCCTCTTCAGTGAAGCTTTCTCCGATTTCCTCATGGAAAATCCACAGATTGCCAAAC
GTATCGTAGAAAAAGGAATTTTGGCTGCCAAGGCTCGTGTGGCTGCCAAGCGTGCGCGTGAAGTCACACGTAAAA
AATCTGGTTTGGAAATTTCCAACCTTCCAGGGAAACTAGCAGACTGTTCTTCTAATAACCCTGCTGAAACAGAACT
CTTCATCGTCGAAGGAGACTCAGCTGGTGGATCAGCCAAATCTGGTCGTAACCGTGAGTTTCAGGCTATCCTTCCA
ATTCGCGGTAAGATTTTGAACGTTGAAAAAGCAAGTATGGATAAGATTCTAGCCAACGAAGAAATTCGTAGTCTTT
TCACAGCCATGGGAACAGGATTTGGCGCAGAATTTGATGTTTCGAAAGCCCGTTACCAAAAACTCGTTTTGATGAC
CGATGCCGATGTCGATGGAGCCCACATTCGTACCCTTCTTTTAACCTTGATTTATCGTTATATGAAACCAATCCTA
GAAGCTGGTTATGTTTATATTGCCCAACCACCAATCTATGGTGTCAAGGTTGGAAGCGAGATTAAAGAATATATCC
AGCCGGGTGCAGATCAAGAAATCAAACTCCAAGAAGCTTTAGCCCGTTATAGTGAAGGTCGTACCAAACCGACTA
TTCAGCGTTATAAGGGGCTAGGTGAAATGGACGATCATCAGCTGTGGGAAACAACCATGGATCCCGAACATCGCT
TGATGGCTAGAGTTTCTGTAGATGATGTGCAGAAGCAGATAAAATCTTTGATATGTTGA

MTEEIKNLQAQDYDASQIQVLEGLEAVRMRPGMYIGSTSKEGLHHLVWEIVDNSIDEALAGFASHIQVFIEPDDSITVVD
DGRGIPVDIQEKTGRPAVETVFTVLHAGGKFGGGGYKVSGGLHGVGSSVVNALSTQLDVHVHKNGKIHYQEYRRGHV
VADLEIVGDTDKTGTTVHFTPDPKIFTETTIFDFDKLNKRIQELAFLNRGLQISITDKRQGLEQTKHYHYEGGIASYVEYI
NENKDVIFDTPIYTDGEMDDITVEVAMQYTTGYHENVMSFANNIHTHEGGTHEQGFRTALTRVINDYARKNKLLKDN
EDNLTGEDVREGLTAVISVKHPNPQFEGQTKTKLGNSEVVKITNRLFSEAFSDFLMENPQIAKRIVEKGILAAKARVAAK
RAREVTRKKSGLEISNLPGKLADCSSNNPAETELFIVEGDSAGGSAKSGRNREFQAILPIRGKILNVEKASMDKILANEEIR
SLFTAMGTGFGAEFDVSKARYQKLVLMTDADVDGAHIRTLLLTLIYRYMKPILEAGYVYIAQPPIYGVKVGSEIKEYIQP
GADQEIKLQEALARYSEGRTKPTIQRYKGLGEMDDHQLWETTMDPEHRLMARVSVDDVQKQIKSLICZ

ID54 1446bp

ATGAGTAGACGTTTTAAAAAATCACGTTCACAGAAAGTGAAGCGAAGTGTTAATATAGTTTTGCTGACTATTTATT
TATTGTTAGTTTGTTTTTTATTGTTCTTAATCTTTAAGTACAATATCCTTGCTTTTAGATATCTTAATCTAGTGGTAA
CTGCGTTAGTCCTACTAGTTGCCTTGGTAGGGCTACTCTTGATTATCTATAAAAAAGCTGAAAAGTTTACTATTTTT
CTGTTGGTGTTCTCTATCCTTGTCAGCTCTGTGTCGCTCTTTGCAGTACAGCAGTTTGTTGGACTGACCAATCGTTT
AAATGCGACTTCTAATTACTCAGAATATTCAATCAGTGTCGCTGTTTTAGCAGATAGTGAGATCGAAAATGTTACG
CAACTGACGAGTGTGACAGCACCGACTGGGACTAATAATGAAAATATTCAGAAATTACTAGCTGATATCAAGTCA
AGTCAGAATACCGATTTGACGGTCAACCAGAGTTCGTGTTACTTGGCAGCTTACAAGAGTTTGATTGCAGGGGAGA
CTAAGGCCATTGTCCTAAATAGTGTCTTTGAAAACATCATCGAGTCAGAGTATCCAGACTACGCATCGAAGATAA
AAAAGATTTATACTAAGGGATTCACTAAAAAAGTAGAAGCTCCTAAGACGTCTAAGAGTCAGTCTTTCAATATCTA
TGTTAGTGGAATTGACACCTATGGTCCTATTAGTTCGGTGTCGCGATCAGATGTCAACATCCTGATGACTGTCAAT
CGAGATACCAAGAAAATCCTCTTGACCACAACGCCACGTGATGCCTATGTACCAATCGCAGATGGTGGAAATAAT
CAAAAAGATAAATTGACTCATGCGGGCATTTATGGAGTTGATTCGTCCATTCACACCTTAGAAAATCTCTATGGAG
TGGATATCAATTACTATGTGCGATTGAACTTCACTTCGTTTTTGAAATTGATTGATTTGTTGGGTGGAATTGATGTT
TATAATGATCAAGAATTTACTGCCCATACGAATGGAAAGTATTACCCTGCAGGCAATGTTCATCTTGATTCAGAAC
AGGCTCTCGGTTTTGTTCGTGAGCGCTACTCCCTAGCAGATGGCGATCGTGACCGCGGGCGCCATCAACAAAAGG
TGATTGTGGCTATCCTTCAAAAATTAACGTCAACCGAAGTGCTGAAAAATTATAGTACGATCATTAATAGCTTGCA
AGATTCTATCCAAACAAATATGCCACTTGAGACCATGATAAATTTGGTCAATGCTCAGTTAGAAAGTGGAGGGAA
TTATAAAGTAAATTCTCAAGATTTAAAAGGGACAGGTCGGATGGATCTTCCTTCTTATGCAATGCCAGACAGTAAC
CTCTATGTGATGGAAATAGATGATAGTAGTTTAGCTGTAGTTAAAGCAGCTATACAGGATGTGATGGAGGGTAGA
TGA

MSRRFKKSRSQKVKRSVNIVLLTIYLLLVCFLLFLIFKYNILAFRYLNLVVTALVLLVALVGLLLIIYKKAEKFTIFLLVFSI
LVSSVSLFAVQQFVGLTNRLNATSNYSEYSISVAVLADSEIENVTQLTSVTAPTGTNNENIQKLLADIKSSQNTDLTVNQS
SSYLAAYKSLIAGETKAIVLNSVFENIIESEYPDYASKIKKIYTKGFTKKVEAPKTSKSQSFNIYVSGIDTYGPISSVSRSDVN
ILMTVNRDTKKILLTTTPRDAYVPIADGGNNQKDKLTHAGIYGVDSSIHTLENLYGVDINYYVRLNFTSFLKLIDLLGGI
DVYNDQEFTAHTNGKYYPAGNVHLDSEQALGFVRERYSLADGDRDRGRHQQKVIVAILQKLTSTEVLKNYSTIINSLQ
DSIQTNMPLETMINLVNAQLESGGNYKVNSQDLKGTGRMDLPSYAMPDSNLYVMEIDDSSLAVVKAAIQDVMEGRZ

ID55 732bp

ATGATAGACATCCATTCGCATATCGTTTTTGATGTAGATGACGGTCCCAAGTCAAGAGAGGAAAGCAAGGCTCTCT
TGGCAGAATCCTACAGACAGGGGGTGCGAACCATTGTTTCTACCTCTCACCGTCGCAAGGGCATGTTTGAAACTCC
GGAAGAGAAGATAGCAGAAAACTTTCTTCAGGTTCGGGAAATAGCTAAGGAAGTGGCGAGTGACTTGGTCATTGC
TTACGGGGCTGAAATTTATTACACACCAGATGTTCTGGATAAGCTGGAAAAAAAGCGGATTCCGACCCTCAATGA
TAGTCGTTATGCCTTGATAGAGTTTAGTATGAACACTCCTTATCGCGATATTCATAGCGCCTTGAGCAAGATCTTG
ATGTTGGGAATTACTCCAGTCATTGCCCACATTGAGCGCTATGATGCTCTTGAAAATAATGAAAAACGCGTTCGAG
AACTGATCGATATGGGCTGTTACACGCAAGTAAATAGTTCACATGTCCTCAAACCCAAACTTTTTGGCGAACGTTA
TAAATTCATGAAAAAAAGAGCTCAGTATTTTTTAGAGCAGGATTTGGTTCATGTCATTGCAAGTGATATGCACAAT
CTAGACGGTAGACCTCCTCATATGGCAGAAGCATATGACCTTGTTACCCAAAAATACGGAGAAGCGAAGGCTCAG
GAACTTTTTATAGACAATCCTCGAAAAATTGTAATGGATCAACTAATTTAG

MIDIHSHIVFDVDDGPKSREESKALLAESYRQGVRTIVSTSHRRKGMFETPEEKIAENFLQVREIAKEVASDLVIAYGAEI
YYTPDVLDKLEKKRIPTLNDSRYALIEFSMNTPYRDIHSALSKILMLGITPVIAHIERYDALENNEKRVRELIDMGCYTQV
NSSHVLKPKLFGERYKFMKKRAQYFLEQDLVHVIASDMHNLDGRPPHMAEAYDLVTQKYGEAKAQELFIDNPRKIVM
DQLIZ

ID58 3990bp

TTGATTTATATAATCGCTATCAATATAACAATGCAATCAGGAGGTTTTGCAATGAAACATGAAAAACAACAGCGTT
TTTCTATTCGTAAATACGCTGTAGGAGCAGCTTCTGTTCTAATTGGATTTGCCTTCCAAGCACAGACTGTTGCAGC
CGATGGAGTTACTCCTACTACTACAGAAAACCAACCGACCATCCATACGGTTTCTGATTCCCCTCAATCATCCGAA
AATCGGACTGAGGAAACACCTAAAGCAGTGCTTCAACCAGAAGCTCCAAAAACTGTAGAAACAGAAACTCCAGCT
ACTGATAAGGTAGCTAGTCTTCCAAAAACAGAAGAAAAACCACAAGAGGAAGTTAGTTCAACTCCTAGTGATAAA
GCAGAAGTGGTAACTCCAACTTCTGCTGAAAAAGAAACTGCTAATAAAAAGGCAGAAGAAGCTAGCCCTAAAAA
GGAAGAAGCGAAAGAGGTTGATTCTAAAGAGTCAAATACAGACAAGACTGACAAGGATAAACCAGCTAAAAAAG
ATGAAGCGAAAGCAGAGGCTGACAAACCGGCAACAGAGGCAGGAAAGGAACGTGCTGCAACTGTAAATGAAAAA
CTAGCGAAAAAGAAAATTGTTTCTATTGATGCTGGACGTAAATATTTCTCACCAGAACAGCTCAAGGAAATCATCG
ATAAAGCGAAACATTATGGCTACACTGATTTACACCTATTAGTCGGAAATGATGGACTCCGTTTCATGTTGGACGA
TATGAGCATCACAGCTAACGGCAAGACCTATGCCAGTGACGATGTCAAACGCGCCATTGAAAAAGGTACAAATGA
TTATTACAACGATCCAAACGGCAATCACTTAACAGAAAGTCAAATGACAGATCTGATTAACTATGCCAAAGATAA
AGGTATCGGTCTCATTCCGACAGTAAATAGTCCTGGACACATGGATGCGATTCTCAATGCCATGAAAGAATTGGG
AATCCAAAACCCTAACTTTAGCTATTTTGGGAAGAAATCAGCCCGTACTGTCGATCTTGACAACGAACAAGCTGTC
GCTTTTACAAAAGCCCTTATCGACAAGTATGCTGCTTATTTCGCGAAAAAGACTGAAATCTTCAACATCGGACTTG
ATGAATATGCCAATGATGCGACAGATGCTAAAGGTTGGAGTGTGCTTCAAGCTGATAAATACTATCCAAACGAAG
GCTACCCTGTAAAAGGCTATGAAAAATTTATTGCCTACGCCAATGACCTCGCTCGTATTGTAAAATCGCACGGTCT
CAAACCAATGGCTTTTAACGACGGTATCTACTACAATAGCGACACAAGCTTTGGTAGTTTTGACAAAGACATCATC
GTTTCTATGTGGACTGGTGGTTGGGGAGGCTACGATGTCGCTTCTTCTAAACTACTAGCTGAAAAAGGTCACCAAA
TCCTTAATACCAATGATGCTTGGTACTACGTTCTTGGACGAAACGCTGATGGCCAAGGCTGGTACAATCTCGATCA
GGGGCTCAATGGTATTAAAAACACACCAATCACTTCTGTACCAAAAACAGAAGGAGCTGATATCCGAATCATCGG
TGGTATGGTAGCTGCTTGGGCTGACACTCCATCTGCACGTTATTCACCATCACGCCTCTTCAAACTCATGCGTCAT
TTTGCAAATGCCAACGCTGAATACTTCGCAGCTGATTATGAATCTGCAGAGCAAGCACTTAACGAGGTACCAAAA
GACCTGAACCGTTATACTGCAGAAAGCGTCACGGCCGTAAAAGAAGCTGAAAAAGCTATTCGCTCTCTCGATAGC
AACCTTAGCCGTGCCCAACAAGATACGATTGATCAAGCCATTGCTAAACTTCAAGAAACTGTCAACAACTTGACC
CTCACGCCTGAAGCTCAAAAAGAAGAAGAAGCTAAACGTGAGGTTGAAAAACTTGCCAAAAACAAGGTAATCTCA
ATCGATGCTGGACGCAAATACTTTACTCTGAACCAGCTCAAACGCATCGTAGACAAGGCCAGTGAGCTCGGATAT
TCTGATGTCCATCTCCTTCTAGGAAATGACGGACTTCGCTTTCTACTCGATGATATGACCATTACTGCCAACGGAA
AAACCTATGCTAGTGATGACGTTAAAAAAGCTATTATCGAAGGAACTAAAGCTTACTACGACGATCCAAACGGTA
CTGCACTAACACAGGCAGAAGTAACAGAGCTAATTGAATACGCTAAATCTAAGGACATCGGTCTCATCCCAGCTA
TTAACAGTCCAGGTCACATGGATGCTATGCTGGTTGCCATGGAAAAATTAGGTATTAAAAATCCTCAAGCCCACTT
TGATAAAGTTTCAAAAACAACTATGGACTTGAAAAACGAAGAAGCGATGAACTTTGTAAAAGCCCTCATCGGTAA
ATACATGGACTTCTTTGCAGGTAAAACAAAGATTTTCAACTTTGGTACTGACGAATACGCCAACGATGCGACTAGT
GCCCAAGGCTGGTACTACCTCAAGTGGTATCAACTCTATGGCAAATTTGCCGAATATGCCAACACCCTCGCAGCTA
TGGCCAAAGAAAGAGGGCTTCAACCAATGGCCTTCAACGATGGCTTCTACTATGAAGACAAGGACGATGTTCAGT
TTGACAAAGATGTCTTGATTTCTTACTGGTCTAAAGGCTGGTGGGGATATAACCTCGCATCACCTCAATACCTAGC
AAGCAAAGGCTATAAATTCTTGAATACCAACGGTGACTGGTACTACATTCTTGGTCAAAAACCAGAAGATGGTGG
TGGTTTCCTCAAGAAAGCTATTGAGAATACTGGAAAAACACCATTCAATCAACTAGCTTCTACCAAATATCCTGAA
GTAGATCTTCCAACAGTCGGAAGTATGCTTTCAATCTGGGCAGATAGACCAAGCGCTGAATACAAGGAAGAGGAA
ATCTTTGAACTCATGACTGCCTTTGCAGACCACAACAAAGACTACTTTCGTGCTAATTATAATGCTCTCCGCGAAG
AATTAGCTAAAATTCCTACAAACTTAGAAGGATATAGTAAAGAAAGTCTTGAGGCCCTTGACGCAGCTAAAACAG
CTCTAAATTACAACCTCAACCGTAATAAACAAGCTGAGCTTGACACGCTTGTAGCCAACCTAAAAGCCGCTCTTCA
AGGCCTCAAACCAGCTGTAACTCATTCAGGAAGCCTAGATGAAAATGAAGTGGCTGCCAATGTTGAAACCAGACC
AGAACTCATCACAAGAACTGAAGAAATTCCATTTGAAGTTATCAAGAAAGAAAATCCTAACCTCCCAGCCGGTCA

GGAAAATATTATCACAGCAGGAGTCAAAGGTGAACGAACTCATTACATCTCTGTACTCACTGAAAATGGAAAAAC
AACAGAAACAGTCCTTGATAGCCAGGTAACCAAAGAAGTTATAAACCAAGTGGTTGAAGTTGGCGCTCCTGTAAC
TCACAAGGGTGATGAAAGTGGTCTTGCACCAACTACTGAGGTAAAACCTAGACTGGATATCCAAGAAGAAGAAAT
TCCATTTACCACAGTGACTTGTGAAAATCCACTCTTACTCAAAGGAAAAACACAAGTCATTACTAAGGGCGTCAAT
GGACATCGTAGCAACTTCTACTCTGTGAGCACTTCTGCCGATGGTAAGGAAGTGAAAACACTTGTAAATAGTGTCG
TAGCACAGGAAGCCGTTACTCAAATAGTCGAAGTCGGAACTATGGTAACACATGTAGGCGATGAAAACGGACAAG
CCGCTATTGCTGAAGAAAAACCAAAACTAGAAATCCCAAGCCAACCAGCTCCATCAACTGCTCCTGCTGAGGAAA
GCAAAGTTCTTCCTCAAGATCCAGCTCCTGTGGTAACAGAGAAAAAACTTCCTGAAACAGGAACTCACGATTCTG
CAGGACTAGTAGTCGCAGGACTCATGTCCACACTAGCAGCCTATGGACTCACTAAAAGAAAAGAAGACTAA

MIYIIAINITMQSGGFAMKHEKQQRFSIRKYAVGAASVLIGFAFQAQTVAADGVTPTTTENQPTIHTVSDSPQSSENRTEE
TPKAVLQPEAPKTVETETPATDKVASLPKTEEKPQEEVSSTPSDKAEVVTPTSAEKETANKKAEEASPKKEEAKEVDSKE
SNTDKTDKDKPAKKDEAKAEADKPATEAGKERAATVNEKLAKKKIVSIDAGRKYFSPEQLKEIIDKAKHYGYTDLHLL
VGNDGLRFMLDDMSITANGKTYASDDVKRAIEKGTNDYYNDPNGNHLTESQMTDLINYAKDKGIGLIPTVNSPGHMD
AILNAMKELGIQNPNFSYFGKKSARTVDLDNEQAVAFTKALIDKYAAYFAKKTEIFNIGLDEYANDATDAKGWSVLQA
DKYYPNEGYPVKGYEKFIAYANDLARIVKSHGLKPMAFNDGIYYNSDTSFGSFDKDIIVSMWTGGWGGYDVASSKLLA
EKGHQILNTNDAWYYVLGRNADGQGWYNLDQGLNGIKNTPITSVPKTEGADIPIIGGMVAAWADTPSARYSPSRLFKL
MRHFANANAEYFAADYESAEQALNEVPKDLNRYTAESVTAVKEAEKAIRSLDSNLSRAQQDTIDQAIAKLQETVNNLT
LTPEAQKEEEAKREVEKLAKNKVISIDAGRKYFTLNQLKRIVDKASELGYSDVHLLLGNDGLRFLLDDMTITANGKTYA
SDDVKKAIIEGTKAYYDDPNGTALTQAEVTELIEYAKSKDIGLIPAINSPGHMDAMLVAMEKLGIKNPQAHFDKVSKTT
MDLKNEEAMNFVKALIGKYMDFFAGKTKIFNFGTDEYANDATSAQGWYYLKWYQLYGKFAEYANTLAAMAKERGL
QPMAFNDGFYYEDKDDVQFDKDVLISYWSKGWWGYNLASPQYLASKGYKFLNTNGDWYYILGQKPEDGGGFLKKAI
ENTGKTPFNQLASTKYPEVDLPTVGSMLSIWADRPSAEYKEEEIFELMTAFADHNKDYFRANYNALREELAKIPTNLEG
YSKESLEALDAAKTALNYNLNRNKQAELDTLVANLKAALQGLKPAVTHSGSLDENEVAANVETRPELITRTEEIPFEVI
KKENPNLPAGQENIITAGVKGERTHYISVLTENGKTTETVLDSQVTKEVINQVVEVGAPVTHKGDESGLAPTTEVKPRL
DIQEEEIPFTTVTCENPLLLKGKTQVITKGVNGHRSNFYSVSTSADGKEVKTLVNSVVAQEAVTQIVEVGTMVTHVGDE
NGQAAIAEEKPKLEIPSQPAPSTAPAEESKVLPQDPAPVVTEKKLPETGTHDSAGLVVAGLMSTLAAYGLTKRKEDZ

ID122 825bp

ATGAACAAAAAAACAAGACAGACACTAATCGGACTGCTAGTGTTATTGCTTTTGTCTACAGGGAGCTATTATATCA
AGCAGATGCCGTCGGCACCTAATAGTCCCAAAACCAATCTTAGTCAGAAAAAACAAGCGTCTGAAGCTCCTAGTC
AAGCATTGGCAGAGAGTGTCTTAACAGACGCAGTCAAGAGTCAAATAAAGGGGAGTCTGGAGTGGAATGGCTCAG
GTGCTTTTATCGTCAATGGTAATAAAACAAATCTAGATGCCAAGGTTTCAAGTAAGCCCTACGCTGACAATAAAAC
AAAGACAGTGGGCAAGGAAACTGTTCCAACCGTAGCTAATGCCCTCTTGTCTAAGGCCACTCGTCAGTACAAGAA
TCGTAAAGAAACTGGGAATGGTTCAACTTCTTGGACTCCTCCAGGTTGGCATCAGGTCAAGAATCTAAAGGGCTCT
TATACCCATGCAGTCGATAGAGGTCATTTGTTAGGCTATGCCTTAATCGGTGGTTTGGATGGTTTTGATGCCTCAA
CAAGCAATCCTAAAAACATTGCTGTTCAGACAGCCTGGGCAAATCAGGCACAAGCCGAGTATTCGACTGGTCAAA
ACTACTATGAAAGCAAGGTGCGTAAAGCCTTGGACCAAAACAAGCGTGTCCGTTACCGTGTAACCCTTTACTACG
CTTCAAACGAGGATTTAGTTCCCTCAGCTTCACAGATTGAAGCCAAGTCTTCGGATGGAGAATTGGAATTCAATGT
TCTAGTTCCCAATGTTCAAAAGGGACTTCAACTGGATTACCGAACTGGAGAAGTAACTGTAACTCAGTAA

MNKKTRQTLIGLLVLLLLSTGSYYIKQMPSAPNSPKTNLSQKKQASEAPSQALAESVLTDAVKSQIKGSLEWNGSGAFIV
NGNKTNLDAKVSSKPYADNKTKTVGKETVPTVANALLSKATRQYKNRKETGNGSTSWTPPGWHQVKNLKGSYTHAV
DRGHLLGYALIGGLDGFDASTSNPKNIAVQTAWANQAQAEYSTGQNYYESKVRKALDQNKRVRYRVTLYYASNEDLV
PSASQIEAKSSDGELEFNVLVPNVQKGLQLDYRTGEVTVTQZ

ID123 225bp

GTGCTAAGATTCAGCGGATTGAGGCAAGTGATGAAGATGAATAAGAAATCAAGCTACGTAGTCAAGCGTTTACTT
TTAGTCATCATAGTACTGATTTTAGGTACTCTGGCTCTAGGAATCGGTTTAATGGTAGGTTATGGAATCTTGGGCA
AGGGTCAAGATCCATGGGCTATCCTGTCTCCAGCAAAATGGCAGGAATTGATTCATAAATTTACAGGAAATTAG

VLRFSGLRQVMKMNKKSSYVVKRLLLVIIVLILGTLALGIGLMVGYGILGKGQDPWAILSPAKWQELIHKFTGNZ

CLAIMS:

1. A *Streptococcus pneumoniae* protein or polypeptide having a sequence selected from those shown in table 1.

2. A *Streptococcus pneumoniae* protein or polypeptide having a sequence selected from those shown in table 2.

3. A protein or polypeptide as claimed in claim 1 or claim 2 provided in substantially pure form.

4. A protein or polypeptide which is substantially identical to one defined in any one of claims 1 to 3.

5. A homologue or derivative of a protein or polypeptide as defined in any one of claims 1 to 4.

6. An antigenic and/or immunogenic fragment of a protein or polypeptide as defined in Tables 1-3.

7. A nucleic acid molecule comprising or consisting of a sequence which is:

(i) any of the DNA sequences set out in Table 1 or their RNA equivalents;

(ii) a sequence which is complementary to any of the sequences of (i);

(iii) a sequence which codes for the same protein or polypeptide, as those sequences of (i) or (ii);

(iv) a sequence which is substantially identical with any of those of (i), (ii) and (iii);

(v) a sequence which codes for a homologue, derivative or fragment of a protein as defined in Table 1.

8. A nucleic acid molecule comprising or consisting of a sequence which is:

(i) any of the DNA sequences set out in Table 2 or their RNA equivalents;

(ii) a sequence which is complementary to any of the sequences of (i);

(iii) a sequence which codes for the same protein or polypeptide, as those sequences of (i) or (ii);

(iv) a sequence which is substantially identical with any of those of (i), (ii) and (iii);

(v) a sequence which codes for a homologue, derivative or fragment of a protein as defined in Table 2.

9. The use of a protein or polypeptide having a sequence selected from those shown in Tables 1-3, or homologues, derivatives and/or fragments thereof, as an immunogen and/or antigen.

10. An immunogenic and/or antigenic composition comprising one or more proteins or polypeptides selected from those whose sequences are shown in Tables 1-3, or homologues or derivatives thereof, and/or fragments of any of these.

11. An immunogenic and/or antigenic composition as claimed in claim 10 which is

a vaccine or is for use in a diagnostic assay.

12. A vaccine as claimed in claim 11 which comprises one or more additional components selected from excipients, diluents, adjuvants or the like.

13. A vaccine composition comprising one or more nucleic acid sequences as defined in Tables 1-3.

14. A method for the detection/diagnosis of *S.pneumoniae* which comprises the step of bringing into contact a sample to be tested with at least one protein or polypeptide as defined in Tables 1-3, or homologue, derivative or fragment thereof.

15. An antibody capable of binding to a protein or polypeptide as defined in Tables 1-3, or for a homologue, derivative or fragment thereof.

16. An antibody as defined in claim 15 which is a monoclonal antibody.

17. A method for the detection/diagnosis of *S.pneumoniae* which comprises the step of bringing into contact a sample to be tested and at least one antibody as define din claim 15 or claim 16.

18. A method for the detection/diagnosis of *S.pneumoniae* which comprises the step of bringing into contact a sample to be tested with at least one nucleic acid sequence as defined in claim 7 or claim 8.

19. A method of determining whether a protein or polypeptide as defined in Tables 1-3 represents a potential anti-microbial target which comprises inactivating said protein or polypeptide and determining whether *S.pneumoniae* is still viable.

20. The use of an agent capable of antagonising, inhibiting or otherwise interfering

with the function or expression of a protein or polypeptide as defined in Tables 1-3 in the manufacture of a medicament for use in the treatment or prophylaxis of *S.pneumoniae* infection